

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:25:39 ; Search time 70.371 Seconds
(without alignments)
1758.725 Million cell updates/sec

Title: US-10-622-817-2

Perfect score: 1667
Sequence: 1 MPWQVRYHGGAPLRYEL.....RMNRCENLAPNTLKLK 320

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq.1Dec04:*

1: geneeqp19808:*\n2: geneeqp19908:*\n3: geneeqp20008:*\n4: geneeqp20018:*\n5: geneeqp20028:*\n6: geneeqp20038:*\n7: geneeqp20038:*\n8: geneeqp20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 1667 | 100.0 | 320 | 7 | AD58613 |
| 2 | 1667 | 100.0 | 320 | 8 | ABM80843 |
| 3 | 1630 | 97.8 | 312 | 2 | AAW25776 |
| 4 | 1630 | 97.8 | 312 | 8 | ADR86551 |
| 5 | 1378 | 82.7 | 272 | 7 | ADC10204 |
| 6 | 1391 | 71.4 | 229 | 8 | ADR86553 |
| 7 | 832 | 48.9 | 161 | 8 | ADR86552 |
| 8 | 247 | 14.8 | 51 | 5 | ABP01502 |
| 9 | 214.5 | 12.9 | 334 | 4 | ABB62468 |
| 10 | 181 | 10.9 | 263 | 4 | ABG23964 |
| 11 | 125.5 | 7.5 | 716 | 3 | AA641411 |
| 12 | 125.5 | 7.5 | 719 | 3 | AA641410 |
| 13 | 125.5 | 7.5 | 748 | 3 | AA641409 |
| 14 | 104 | 6.2 | 639 | 4 | ABB62218 |
| 15 | 103 | 6.2 | 1512 | 4 | AAU04349 |
| 16 | 103 | 6.2 | 1512 | 4 | AAW78732 |
| 17 | 103 | 6.2 | 1550 | 4 | AAW79716 |
| 18 | 102 | 6.1 | 2001 | 8 | AD524684 |
| 19 | 101 | 6.1 | 559 | 2 | AA10682 |
| 20 | 101 | 6.1 | 826 | 4 | AA879740 |
| 21 | 101 | 6.1 | 826 | 4 | AA879989 |
| 22 | 101 | 6.1 | 826 | 4 | AAU71908 |
| 23 | 101 | 6.1 | 833 | 4 | AA879739 |
| 24 | 101 | 6.1 | 833 | 4 | AA879988 |
| 25 | 101 | 6.1 | 833 | 4 | AAU71907 |

| | | | | | | |
|----|-------|-----|------|---|----------|--------------------|
| 26 | 101 | 6.1 | 1221 | 4 | AA691409 | AA691409 C glutam |
| 27 | 101 | 6.1 | 1221 | 5 | ABG80321 | ABG80321 C. glutam |
| 28 | 101 | 6.1 | 1221 | 5 | AAE18908 | AAE18908 Coryneb |
| 29 | 101 | 6.1 | 1221 | 7 | ADD13589 | ADD13589 C. glutam |
| 30 | 100.5 | 6.0 | 1200 | 6 | ABU26069 | ABU26069 Protein E |
| 31 | 99.5 | 6.0 | 650 | 2 | AA729621 | AA729621 Mouse P1A |
| 32 | 97.5 | 5.8 | 651 | 2 | AA729621 | AA729621 Mouse P1A |
| 33 | 97.5 | 5.8 | 217 | 7 | AD63092 | AD63092 Rat Prote |
| 34 | 97.5 | 5.8 | 217 | 7 | ADD48150 | ADD48150 Rat Prote |
| 35 | 97.5 | 5.8 | 217 | 7 | ADD48158 | ADD48158 Rat Prote |
| 36 | 97.5 | 5.8 | 217 | 7 | ADD48154 | ADD48154 Rat Prote |
| 37 | 97.5 | 5.8 | 217 | 7 | AD63088 | AD63088 Rat Prote |
| 38 | 97.5 | 5.8 | 217 | 7 | AD63088 | AD63088 Rat Prote |
| 39 | 97.5 | 5.8 | 218 | 2 | AA20033 | AA20033 Rat brain |
| 40 | 97.5 | 5.8 | 218 | 2 | AD56502 | AD56502 Rat Prote |
| 41 | 97.5 | 5.8 | 652 | 5 | ABP41869 | ABP41869 Human ova |
| 42 | 97 | 5.8 | 559 | 3 | AA77178 | AA77178 S. venez |
| 43 | 97 | 5.8 | 2221 | 4 | ABB61339 | ABB61339 Drosophi |
| 44 | 97 | 5.8 | 5215 | 8 | ADL91930 | ADL91930 Streptomy |
| 45 | 96.5 | 5.8 | 1049 | 6 | ABM67090 | ABM67090 Photorhab |

ALIGNMENTS

| | |
|----------|---|
| RESULT 1 | |
| AD58613 | standard; protein; 320 AA. |
| ID | AD58613 |
| AC | AD58613; |
| XX | |
| DT | 29-JAN-2004 (first entry) |
| XX | |
| DE | Human Protein Q13155, SEQ ID NO 4489. |
| XX | |
| KW | Human; pain; neuronal tissue; gene therapy; |
| KW | spinal segmental nerve injury; chronic constriction injury; CCI; |
| KW | spared nerve injury; SNF; Chung. |
| OS | Homo sapiens. |
| PN | WO2003016475-A2. |
| XX | |
| PD | 27-FEB-2003. |
| XX | |
| PF | 14-AUG-2002; 2002WO-US025765. |
| XX | |
| PR | 14-AUG-2001; 2001US-0312147P. |
| PR | 01-NOV-2001; 2001US-0346382P. |
| PR | 26-NOV-2001; 2001US-0333347P. |
| XX | |
| PA | (GENO) GEN HOSPITAL CORP. |
| XX | (FARB) BAYER AG. |
| XX | |
| PI | Wolff C, D'urso D, Bofort K, Costigan M; |
| XX | |
| DR | WPI; 2003-268312/26. |
| DR | GENBANK; Q13155. |
| XX | |
| PT | New composition comprising two or more isolated polypeptides, useful for |
| PT | preparing a medicament for treating pain in an animal. |
| XX | |
| PS | Claim 1; Page; 1017pp; English. |
| XX | |
| CC | The invention discloses a composition comprising two or more isolated rat |
| CC | or human polynucleotides or a polynucleotide which represents a fragment, |
| CC | derivative or allelic variation of the nucleic acid sequence. Also |
| CC | claimed are a vector comprising the novel polynucleotide, a host cell |
| CC | comprising the vector, a method for identifying a nucleotide sequence |
| CC | which is differentially regulated in an animal subjected to pain and a |
| CC | kit to perform the method, an array, a method for identifying an agent |
| CC | that increases or decreases the expression of the polynucleotide sequence |
| CC | that is differentially expressed in neuronal tissue of a first animal |

CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition,
CC a method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (thung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC the sequence data for this patent did not form part of the printed
CC ftp.wipo.int/pub/published_pct_sequences.

Sequence 320 AA;

Query Match
Best Local Similarity 100.0%; Score 1667; DB 7; Length 320;
Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPMTQVRYHGGAPLREVELPTCMYRLPMVHGRSYGAPGAGHVOESNLSIQALSROD 60
DB 1 MPMTQVRYHGGAPLREVELPTCMYRLPMVHGRSYGAPGAGHVOESNLSIQALSROD 60
QY 61 DIKRLYLEKKAANDGSKMIOPTPDADLVNIIQADEPTTLTNALDLSVGYGKYGALK 120
DB 61 DIKRLYLEKKAANDGSKMIOPTPDADLVNIIQADEPTTLTNALDLSVGYGKYGALK 120
QY 121 DIVYNNAPSPSLSLVYHRLCEHFRVLSTVHTHSVSVBENLKCFGEONKKOPROD 180
DB 121 DIVYNNAPSPSLSLVYHRLCEHFRVLSTVHTHSVSVBENLKCFGEONKKOPROD 180
QY 181 YOLGTTLIWKVPTQMKFSIQMCEPIBEGNINARFSLFGQKHNANATLIDSWDIA 240
DB 181 YOLGTTLIWKVPTQMKFSIQMCEPIBEGNINARFSLFGQKHNANATLIDSWDIA 240
QY 241 IFOLKESSKEKAAVPRSNMAGKSPMLAGNELTVADVVMSTLQOIGGCVTVVPAVNO 300
DB 241 IFOLKESSKEKAAVPRSNMAGKSPMLAGNELTVADVVMSTLQOIGGCVTVVPAVNO 300
QY 301 RMRSCENLAPFNTALKLK 320
DB 301 RMRSCENLAPFNTALKLK 320

RESULT 2

ABM80843
ID ABM80843 standard; protein; 320 AA.
XX
AC ABM80843;
XX
XX 18-NOV-2004 (first entry)

DE Tumour-associated antigenic target (TAT) polypeptide PRO81501, SEQ:2179.
XX
XX Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW cervical cancer; melanoma; leukemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic.

OS Homo sapiens.
XX
XX MO2004030615-A2.
XX
XX 15-APR-2004.

XX
PF 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-041971P.
XX
PA (GETH) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;
XX
DR MPI; 2004-347921/32.
XX
DR N-PSDB; ACN36579.
XX
PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.

Claim 12; SEQ ID NO 2179; 7273bp; English.

CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acid and polypeptide
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT polypeptide of the invention

Sequence 320 AA;

Query Match
Best Local Similarity 100.0%; Score 1667; DB 6; Length 320;
Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPMTQVRYHGGAPLREVELPTCMYRLPMVHGRSYGAPGAGHVOESNLSIQALSROD 60
DB 1 MPMTQVRYHGGAPLREVELPTCMYRLPMVHGRSYGAPGAGHVOESNLSIQALSROD 60
QY 61 DIKRLYLEKKAANDGSKMIOPTPDADLVNIIQADEPTTLTNALDLSVGYGKYGALK 120
DB 61 DIKRLYLEKKAANDGSKMIOPTPDADLVNIIQADEPTTLTNALDLSVGYGKYGALK 120
QY 121 DIVYNNAPSPSLSLVYHRLCEHFRVLSTVHTHSVSVBENLKCFGEONKKOPROD 180
DB 121 DIVYNNAPSPSLSLVYHRLCEHFRVLSTVHTHSVSVBENLKCFGEONKKOPROD 180
QY 181 YOLGTTLIWKVPTQMKFSIQMCEPIBEGNINARFSLFGQKHNANATLIDSWDIA 240
DB 181 YOLGTTLIWKVPTQMKFSIQMCEPIBEGNINARFSLFGQKHNANATLIDSWDIA 240
QY 241 IFOLKESSKEKAAVPRSNMAGKSPMLAGNELTVADVVMSTLQOIGGCVTVVPAVNO 300
DB 241 IFOLKESSKEKAAVPRSNMAGKSPMLAGNELTVADVVMSTLQOIGGCVTVVPAVNO 300
QY 301 RMRSCENLAPFNTALKLK 320
DB 301 RMRSCENLAPFNTALKLK 320

RESULT 3
AAW25776

ID AAW25776 standard; protein; 312 AA.
XX
AC AAW25776;
XX
DT 19-DEC-1997 (first entry)
XX
DE JTV1 protein.
XX
KM JTV1; hPMS2; probe; detection; chromosome 7; deletion;
KM mismatch repair gene; hereditary non-polyposis colorectal cancer;
KM homologous recombination.
XX
OS Homo sapiens.
XX
XX MO9708312-A1.
XX
PD 06-MAR-1997.
XX
PF 26-AUG-1996; 96MO-US013598.
XX
PR 24-AUG-1995; 95US-00518862.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Vogelstein B, Kinzler KW, Nicolaides NC;
XX
DR WPI; 1997-179269/16.
DR N-PSDB; AAT86182.
XX
PT Novel chromosome 7 gene, JTV1 - used for detecting chromosome 7
PT deletions, and PMS2 promoter activity.
XX
PS Claim 5; Fig 2; 31pp; English.
XX
CC This sequence is JTV1 protein and is encoded by DNA isolated from human
CC chromosome 7. The JTV1 coding sequence is located upstream from hPMS2.
CC JTV1 cDNA can be used as probes to detect chromosome 7 deletions.
CC involving JTV1. Due to the overlapping promoter regions, deletions of
CC JTV1 would also affect PMS2 (a mismatch repair gene) expression, leading
CC to hereditary non-polyposis colorectal cancer. JTV1 can also be used to
CC assay activity or competence of the PMS2 promoter region, the presence of
CC JTV1 suggesting that the PMS2 promoter is intact. JTV1 sequences can also
CC be used to guide homologous recombination at the PMS2 locus
XX
XX
SQ Sequence 312 AA;
XX
Query Match 97.8%; Score 1630; DB 2; Length 312;
Best Local Similarity 100.0%; Pred. No. 4.3e-163;
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MEMYQVKPYHGGAPLRVELPTCMYRLPVNHGRSYGPAGAGHVOEESNLSQLAESRD 60
DB 1 MEMYQVKPYHGGAPLRVELPTCMYRLPVNHGRSYGPAGAGHVOEESNLSQLAESRD 60
QY 61 DILKRLYEKAAVDSKMIQTPDADLVNTIIQADEPTTLTTNALDLSVKGKYGALK 120
DB 61 DILKRLYEKAAVDSKMIQTPDADLVNTIIQADEPTTLTTNALDLSVKGKYGALK 120
QY 121 DIVINANPASPPLSLVLRLLCEHFRVLSVHTHSSVKSVPENLLKCFGEQNKQPRD 180
DB 121 DIVINANPASPPLSLVLRLLCEHFRVLSVHTHSSVKSVPENLLKCFGEQNKQPRD 180
QY 181 YOLGFTLIMKNVPTKTMKFSIQMCPISBEGNARFLFSIFGQKNAVNATLIDSVADIA 240
DB 181 YOLGFTLIMKNVPTKTMKFSIQMCPISBEGNARFLFSIFGQKNAVNATLIDSVADIA 240
QY 241 ITOLKSGSSKEKAARVRSNMSALGKSPMLAGNELTVADVVLMSVLAQIIGCSVTVPANQ 300
DB 241 ITOLKSGSSKEKAARVRSNMSALGKSPMLAGNELTVADVVLMSVLAQIIGCSVTVPANQ 300
QY 301 RWRSCENLAPF 312
DB 301 RWRSCENLAPF 312

RESULT 4
ID ADR86551
AC ADR86551;
XX
DT 18-NOV-2004 (first entry)
XX
DE 1-312 amino acid sequence of p38/JTV-1 protein.
XX
KM p38/JTV-1; Cytostatic; cancer; leukemia; anticancer.
XX
OS Homo sapiens.
XX
XX EP1454628-A2.
XX
PD 08-SEP-2004.
XX
PF 09-SEP-2003; 2003EP-00020344.
XX
PR 03-MAR-2003; 2003KR-00013058.
XX
PA (UYSE-) UNIV SEOUL NAT IND FOUND.
XX
PI Kim S, Park B;
XX
DR WPI; 2004-627822/61.
DR N-PSDB; ADR86548.
XX
PT New isolated p38/JTV-1 protein, useful as medicament for treating cancer
PT e.g., Breast, liver, blood, bone, pancreatic, skin, head or neck cancer
PT and cutaneous or intraocular melanoma, as well as for screening new
XX anticancer agents.
XX
PS Claim 5; SEQ ID NO 4; 47pp; English.
XX
CC The present invention relates to an isolated p38/JTV-1 protein for use as
CC medicament. The p38/JTV-1 protein or the pharmaceutical composition is
CC useful as medicament for treating breast cancer, large intestinal cancer,
CC lung cancer, small cell lung cancer, stomach cancer, liver cancer, blood
CC cancer, bone cancer, pancreatic cancer, skin cancer, head or neck cancer,
CC cutaneous or intraocular melanoma, uterine sarcoma, ovarian cancer,
CC rectal cancer, anal cancer, colon cancer, fallopian tube carcinoma,
CC endometrial carcinoma, cervical cancer, vulval cancer, vaginal carcinoma,
CC Hodgkin's disease, esophageal cancer, small intestine cancer, endocrine
CC cancer, thyroid cancer, parathyroid cancer, adrenal cancer, soft tissue
CC tumour, urethral cancer, penile cancer, prostate cancer, chronic or acute
CC leukemia, lymphocytic lymphoma, bladder cancer, kidney cancer, ureter
CC cancer, renal cell carcinoma, renal pelvic carcinoma, CNS tumour, primary
CC CNS lymphoma, bone marrow tumour, brain stem nerve glioma, pituitary
CC adenoma, or their combination. The protein is useful as a target for
CC screening new anticancer agents. The present sequence represents the 1-
CC 312 amino acid sequence of p38/JTV-1 protein.
XX
XX
SQ Sequence 312 AA;
XX
Query Match 97.8%; Score 1630; DB 8; Length 312;
Best Local Similarity 100.0%; Pred. No. 4.3e-163;
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MEMYQVKPYHGGAPLRVELPTCMYRLPVNHGRSYGPAGAGHVOEESNLSQLAESRD 60
DB 1 MEMYQVKPYHGGAPLRVELPTCMYRLPVNHGRSYGPAGAGHVOEESNLSQLAESRD 60
QY 61 DILKRLYEKAAVDSKMIQTPDADLVNTIIQADEPTTLTTNALDLSVKGKYGALK 120
DB 61 DILKRLYEKAAVDSKMIQTPDADLVNTIIQADEPTTLTTNALDLSVKGKYGALK 120
QY 121 DIVINANPASPPLSLVLRLLCEHFRVLSVHTHSSVKSVPENLLKCFGEQNKQPRD 180
DB 121 DIVINANPASPPLSLVLRLLCEHFRVLSVHTHSSVKSVPENLLKCFGEQNKQPRD 180

```

QY 161 YOLGFTLIWKVPTOMKPSIOTMCPRIEGBGNIAFLFSFGOKHNAVNTLIDSWDIA 240
DB 161 YOLGFTLIWKVPTOMKPSIOTMCPRIEGBGNIAFLFSFGOKHNAVNTLIDSWDIA 240
QY 241 IFOLKESSSKEKAIVFRSMNSALGKSPWLAGNELTVADVLMSTVLTQOIGGCVTPANVQ 300
DB 241 IFOLKESSSKEKAIVFRSMNSALGKSPWLAGNELTVADVLMSTVLTQOIGGCVTPANVQ 300
QY 301 RMRSCENLAPF 312
DB 301 RMRSCENLAPF 312

RESULT 5
ADCI0204
ID ADCI0204 standard; protein; 272 AA.
AC ADCI0204;
XX
XX
XX 18-DEC-2003 (first entry)
XX
XX Human NOVX polypeptide SEQ ID NO: 226.
XX
XX cyrostatic; antidiabetic; anorectic; cerebroprotective; neuroprotective;
XX antiinflammatory; gene therapy; antisense therapy; thymimetic; NOVX;
XX pathology; cancer; diabetes; obesity; endocrine disorder; CNS disorder;
XX inflammatory disorder; chromosome mapping; tissue typing;
XX predictive medicine.
XX
XX Homo sapiens.
XX
XX WO2003000842-A2.
XX
XX 03-JAN-2003.
XX
XX
XX 04-JUN-2002; 2002MO-US017443.
XX
XX 04-JUN-2001; 2001US-0295607P.
XX 04-JUN-2001; 2001US-0295661P.
XX 06-JUN-2001; 2001US-0296404P.
XX 06-JUN-2001; 2001US-0296418P.
XX 07-JUN-2001; 2001US-0296575P.
XX 11-JUN-2001; 2001US-0297414P.
XX 12-JUN-2001; 2001US-0295573P.
XX 12-JUN-2001; 2001US-0297567P.
XX 14-JUN-2001; 2001US-0298285P.
XX 15-JUN-2001; 2001US-0298528P.
XX 18-JUN-2001; 2001US-0299133P.
XX 19-JUN-2001; 2001US-0299230P.
XX 21-JUN-2001; 2001US-0299949P.
XX 22-JUN-2001; 2001US-0300177P.
XX 26-JUN-2001; 2001US-0300883P.
XX 28-JUN-2001; 2001US-0301530P.
XX 28-JUN-2001; 2001US-0301550P.
XX 03-JUL-2001; 2001US-0302951P.
XX 31-JUL-2001; 2001US-0308909P.
XX 14-SEP-2001; 2001US-0322297P.
XX 25-SEP-2001; 2001US-0324699P.
XX 03-DEC-2001; 2001US-0337477P.
XX 14-DEC-2001; 2001US-0341562P.
XX 21-FEB-2002; 2002US-0358656P.
XX 21-FEB-2002; 2002US-0359122P.
XX 22-FEB-2002; 2002US-0358978P.
XX 22-FEB-2002; 2002US-0359034P.
XX 22-FEB-2002; 2002US-0359035P.
XX 27-FEB-2002; 2002US-0359121P.
XX 01-MAR-2002; 2002US-0359684P.
XX 12-MAR-2002; 2002US-0363430P.
XX 12-MAR-2002; 2002US-0363676P.
XX 10-APR-2002; 2002US-0371346P.
XX 10-MAY-2002; 2002US-0379444P.

```

```

PR 04-JUN-2002; 2002US-00379444.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Agee ML, Anderson DM, Berghs C, Casman SJ, Catterton E;
XX diDippo VA, Edinger SR, Eisen A, Ellemann K, Gangoli EA;
XX Gerlach VL, Gorman U, Guo X, Hermann UT, Hjal T, Ji W, Kekuda R;
XX Khrantsov NV, Li L, Liu X, Malyankar UM, Miller CE, Miller I;
XX Ott T, Padigaru M, Patturajan M, Pena CE, Rastelli L, Rieger DK;
XX Rothenberg ME, Shenoy SG, Shinkets RA, Smithson G, Spaderna SK;
XX Spytek KA, Stone DJ, Vernet CAM, Zhong H, Zhong M, Alsbrook UT;
XX Burgess CE, Lepley DM;
XX
XX MPI, 2003-210149/20.
XX
XX N-PSDB; ADCI0203.
XX
XX PT New isolated NOVX polypeptides and nucleic acid molecules useful for
XX PT treating, preventing and diagnosing pathological conditions with NOVX-
XX PT associated disorders, such as cancer, obesity, diabetes and inflammatory
XX PT or CNS diseases.
XX
XX Claim 1; SEQ ID NO 226; 772bp; English.
XX
XX
XX The invention relates to novel isolated polypeptides, mature form of the
XX CC polypeptide, a sequence that is 95% identical to the polypeptide or the
XX CC polypeptide comprising one or more conservative substitutions. The NOVX
XX CC polypeptide is useful for treating or preventing a pathology associated
XX CC with the polypeptide e.g. disorders associated with aberrant expression
XX CC or activity of the polypeptide, such as cancer, diabetes, obesity, and
XX CC endocrine, CNS and inflammatory disorders. They can also be used in
XX CC various detection and screening assays, chromosome mapping, tissue typing
XX CC and predictive medicine. This sequence corresponds to one of the
XX CC polypeptides of the invention.
XX
XX
XX Sequence 272 AA;
XX
XX
XX Query Match 82.7%; Score 1378; DB 7; Length 272;
XX Best Local Similarity 100.0%; Pred. No. 1, 6e-136;
XX Matches 268; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX
XX 45 QEESNLSLOAESRQDDILKRYELKAAVDGSKMIQTDPADVDVNTIIQADEPTTLTN 104
XX |||||
XX 5 QEESNLSLOAESRQDDILKRYELKAAVDGSKMIQTDPADVDVNTIIQADEPTTLTN 64
XX
XX 105 ALDINSYLGKDYGLKXIVINANPASPPLSLVLAHLLCEHFVLSVTWTHSSVKSPEN 164
XX |||||
XX 65 ALDINSYLGKDYGLKXIVINANPASPPLSLVLAHLLCEHFVLSVTWTHSSVKSPEN 124
XX
XX 165 LKCFGEQNKQKPRQDYOLGFTLIWKVPTOMKPSIOTMCPRIEGBGNIAFLFSFGOK 224
XX |||||
XX 125 LKCFGEQNKQKPRQDYOLGFTLIWKVPTOMKPSIOTMCPRIEGBGNIAFLFSFGOK 184
XX
XX QY 225 HNAVNTLIDSWDIAIFOLKESSSKEKAIVFRSMNSALGKSPWLAGNELTVADVLMSTV 284
XX |||||
XX DB 185 HNAVNTLIDSWDIAIFOLKESSSKEKAIVFRSMNSALGKSPWLAGNELTVADVLMSTV 244
XX |||||
XX QY 285 LQOIGGCVTPANVQRMRSCEMLAPF 312
XX |||||
XX DB 245 LQOIGGCVTPANVQRMRSCEMLAPF 272
XX |||||

RESULT 6
ADR86553
ID ADR86553 standard; protein; 229 AA.
XX
XX ADR86553;
XX
XX AC ADR86553;
XX
XX 18-NOV-2004 (first entry)
XX
XX 84-312 amino acid sequence of p38/JTV-1 protein.
XX
XX p38/JTV-1; Cytostatic; cancer; leukemia; anticancer.
XX

```


DE Human ORFX protein sequence SEQ ID NO:2986.
XX
XX
KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumor; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW myasthenia gravis.
XX
XX Homo sapiens.
OS
XX
XX MO200192523-A2.
XX
XX PD 06-DEC-2001.
XX
XX PF 29-MAY-2001; 2001WO-US010836.
XX
XX PR 30-MAY-2000; 2000US-0206132P.
XX
XX PR 29-AUG-2000; 2000US-0228716P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX PI Shinkets RA, Leach MD;
XX
XX WPI: 2002-106308/14.
XX
XX DR N-PSDB; ABLN17254.
XX
XX PT Novel human polypeptides and polynucleotides useful for diagnosing,
XX preventing and treating cardiovascular disease, neurodegenerative,
XX hyperproliferative disorders and autoimmune disorders.
XX
XX PS Disclosure; SEQ ID NO 2986; 1037bp; English.
XX
XX CC The present invention describes substantially purified human proteins
XX (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
XX in the specification). ABLN15762 to ABLN27252 encode the human ORFX
XX proteins given in ABLN00010 to ABLN11500. ORFX proteins are useful for
XX treating or preventing a pathology associated with an ORFX-associated
XX disorder in humans, and in the manufacture of a medicament for treating a
XX syndrome associated with ORFX-associated disorder. ORFX polynucleotide
XX sequences can be used in gene therapy. ORFX sequences can be used in the
XX treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
XX psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
XX osteoarthritis, neurodegenerative disorders, diabetes mellitus, systemic
XX lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
XX storage disease, various immune deficiencies and disorders, infectious
XX diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
XX arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
XX disease and autoimmune inflammatory eye disease. ORFX proteins are also
XX useful for treating burns, incisions, ulcers, for treating osteoporosis,
XX bone degenerative disorders, or periodontal disease, and for gut
XX protection or regeneration and treatment of lung or liver fibrosis,
XX retransfusion injury in various tissues and conditions resulting from
XX systemic cytokine damage. N.B. The sequence data for this patent did not
XX form part of the printed specification, but was obtained in electronic
XX format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 51 AA;
XX
XX Query Match 14.8%; Score 247; DB 5; Length 51;
XX Best Local Similarity 96.1%; Pred. No. 4.5e-18;
XX Matches 49; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 74 DGLSMKIQTPADLDVNNIIADDEPTLTNNALDLSVGLKGYGALKDVI 124
XX DB 1 DGLSMKIQTPADLDVNNIIADDEPTLTNNALDLSVGLKGYGALKDVI 51

ID ABB62468 standard; protein; 334 AA.
XX
XX ABB62468;
XX AC
XX 26-MAR-2002 (first entry)
XX DT
XX DE Drosophila melanogaster polypeptide SEQ ID NO 14196.
XX
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
XX OS Drosophila melanogaster.
XX
XX XX MO200171042-A2.
XX
XX XX PD 27-SEP-2001.
XX
XX XX PF 23-MAR-2001; 2001WO-US009231.
XX
XX XX PR 23-MAR-2000; 2000US-0191637P.
XX
XX XX PR 11-JUL-2000; 2000US-00614150.
XX
XX DA (PEKE) PE CORP NY.
XX
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX DR WPI: 2001-656860/75.
XX
XX DR N-PSDB; ABL06571.
XX
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX
XX PS Disclosure; SEQ ID NO 14196; 21bp + Sequence Listing; English.
XX
XX CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 334 AA;
XX
XX Query Match 12.9%; Score 214.5; DB 4; Length 334;
XX Best Local Similarity 25.6%; Pred. No. 2.8e-13;
XX Matches 85; Conservative 50; Mismatches 120; Indels 77; Gaps 10;
XX
XX QY 18 VELPTCYRLPNV-----HRSYGPAPGAGHVOESNLSIQ----- 54
XX DB 13 IKLPTCMYPLKAVNSLAADSLASGTSASTSCKLEANNRIDRGRMAATCALDLS 72
XX QY 55 -----LESRODILKRYELKAAVDGSKAIQTPDDLDVNTIQADEPT 99
XX DB 73 LGRQIORLKDPTASVAARQEVKQLELKAQLOQR-----AGLVCQ--KTFQHT 123
XX QY 100 TLTNNALDLSVGLKGYGALKDVIYNNAPSPSLILVLRLLCEHFRVLSVPHSSSVK 159
XX DB 124 TAFQNG-----GLKEVPLQDVVINGHPFIPYALLAKNMRNLVITIDVTFTHSTNA 176
XX QY 160 SY-----PENLAKCFGEONKQKOPRODYQLGFTLIMKRVPTQMKFSIQMCPLEGGN 212
XX DB 177 DIGPAAEFENLAKV--PVNPALPK-----ISVILMKNCEHTEMTSSPTVYVYIGEVN 230
XX QY 213 IARFLFSIFGOKNANVAVT---ISWVDIAIFOLKEGSSKEKAVFSSNNSALGKSPWL 269
XX DB 231 IIRYLAGRVPAEYRSGSPLCNEIDLVLDICYQLLRCTHRTQVAMVRLDLKRLQKQYF 290
XX QY 270 AGNELTVADVYVMSVLOQIGGCSVTVPANVOR 301

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:36:20 ; Search time 17.81 Seconds
(without alignments)
1341.256 Million cell updates/sec

Title: US-10-622-817-2

Perfect score: 1667

Sequence: 1 MPWQVKRYHGGAGLRLVEL.....RWRSCENLAPFNTALKLK 320

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 1667 | 100.0 | 341 | 4 | US-09-949-016-11312 |
| 2 | 1630 | 97.8 | 312 | 2 | US-08-518-862C-2 |
| 3 | 103 | 6.2 | 1512 | 3 | US-09-443-184-48 |
| 4 | 97 | 5.8 | 559 | 2 | US-08-756-317-10 |
| 5 | 97 | 5.8 | 559 | 2 | US-09-091-609-4 |
| 6 | 97 | 5.8 | 5215 | 3 | US-09-105-537-2 |
| 7 | 95 | 5.7 | 559 | 4 | US-09-821-016-1 |
| 8 | 95 | 5.7 | 559 | 4 | US-10-259-633-1 |
| 9 | 95 | 5.7 | 559 | 4 | US-10-266-787-1 |
| 10 | 92 | 5.5 | 2954 | 4 | US-09-150-867-1 |
| 11 | 91 | 5.5 | 474 | 4 | US-09-248-796A-20321 |
| 12 | 90.5 | 5.4 | 359 | 3 | US-09-540-824-2 |
| 13 | 90.5 | 5.4 | 597 | 1 | US-08-399-696-102 |
| 14 | 88.5 | 5.3 | 1007 | 4 | US-09-538-092-736 |
| 15 | 88 | 5.3 | 427 | 4 | US-09-949-016-11178 |
| 16 | 88 | 5.3 | 724 | 4 | US-09-900-920-62 |
| 17 | 88 | 5.3 | 883 | 4 | US-09-976-239-2 |
| 18 | 88 | 5.3 | 883 | 4 | US-10-289-779B-2 |
| 19 | 88 | 5.3 | 914 | 4 | US-09-976-239-4 |
| 20 | 88 | 5.3 | 914 | 4 | US-10-289-779B-4 |
| 21 | 86.5 | 5.2 | 761 | 3 | US-09-012-710-13 |
| 22 | 86.5 | 5.2 | 761 | 3 | US-09-556-273-13 |
| 23 | 86 | 5.2 | 211 | 4 | US-09-328-352-6834 |
| 24 | 86 | 5.2 | 351 | 4 | US-08-178-257-6 |
| 25 | 86 | 5.2 | 471 | 3 | US-09-134-001C-4904 |
| 26 | 86 | 5.2 | 1365 | 3 | US-09-376-330-18 |
| 27 | 86 | 5.2 | 1365 | 6 | 5194600-4 |

| | | | | | | |
|----|------|-----|------|---|----------------------|--------------------|
| 28 | 86 | 5.2 | 1365 | 6 | 5194600-4 | Patent No. 5194600 |
| 29 | 85.5 | 5.1 | 1440 | 3 | US-09-357-251-37 | Sequence 37, Appl |
| 30 | 85 | 5.1 | 559 | 4 | US-10-253-509-1 | Sequence 1, Appl |
| 31 | 85 | 5.1 | 1135 | 2 | US-08-469-537A-97 | Sequence 97, Appl |
| 32 | 84.5 | 5.1 | 579 | 4 | US-09-107-433-4488 | Sequence 4488, Ap |
| 33 | 84.5 | 5.1 | 829 | 1 | US-07-670-611-2 | Sequence 2, Appl |
| 34 | 84.5 | 5.1 | 829 | 1 | US-08-220-674-2 | Sequence 2, Appl |
| 35 | 84.5 | 5.1 | 829 | 1 | US-08-445-186-2 | Sequence 2, Appl |
| 36 | 84.5 | 5.1 | 829 | 1 | US-08-446-548-2 | Sequence 2, Appl |
| 37 | 84.5 | 5.1 | 829 | 2 | US-08-446-550-2 | Sequence 2, Appl |
| 38 | 83.5 | 5.0 | 778 | 4 | US-09-583-110-3830 | Sequence 3930, Ap |
| 39 | 83.5 | 5.0 | 786 | 4 | US-09-107-433-3893 | Sequence 3893, Ap |
| 40 | 83 | 5.0 | 559 | 2 | US-09-052-339-1 | Sequence 1, Appl |
| 41 | 83 | 5.0 | 559 | 3 | US-09-385-742B-1 | Sequence 1, Appl |
| 42 | 83 | 5.0 | 559 | 4 | US-09-989-786-1 | Sequence 1, Appl |
| 43 | 83 | 5.0 | 767 | 3 | US-08-836-567-8 | Sequence 8, Appl |
| 44 | 83 | 5.0 | 767 | 4 | US-09-606-304-8 | Sequence 8, Appl |
| 45 | 82.5 | 4.9 | 222 | 4 | US-09-248-796A-18146 | Sequence 18146, A |

ALIGNMENTS

| | | | | | | | | | |
|--|-----|--|-----|--|--|--|--|--|--|
| RESULT 1 | | | | | | | | | |
| US-09-949-016-11312 | | | | | | | | | |
| Sequence 11312, Application US/09949016 | | | | | | | | | |
| Patent No. 6812339 | | | | | | | | | |
| GENERAL INFORMATION: | | | | | | | | | |
| APPLICANT: VENTER, J. Craig et al. | | | | | | | | | |
| TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED | | | | | | | | | |
| FILE REFERENCE: C1001107 | | | | | | | | | |
| CURRENT APPLICATION NUMBER: US/09/949,016 | | | | | | | | | |
| CURRENT FILING DATE: 2000-04-14 | | | | | | | | | |
| PRIOR APPLICATION NUMBER: 60/241,755 | | | | | | | | | |
| PRIOR FILING DATE: 2000-10-20 | | | | | | | | | |
| PRIOR APPLICATION NUMBER: 60/237,768 | | | | | | | | | |
| PRIOR FILING DATE: 2000-10-03 | | | | | | | | | |
| PRIOR APPLICATION NUMBER: 60/231,498 | | | | | | | | | |
| PRIOR FILING DATE: 2000-09-08 | | | | | | | | | |
| NUMBER OF SEQ ID NOS: 207012 | | | | | | | | | |
| SOFTWARE: FASTSEQ for Windows Version 4.0 | | | | | | | | | |
| SEQ ID NO 11312 | | | | | | | | | |
| LENGTH: 341 | | | | | | | | | |
| TYPE: PRT | | | | | | | | | |
| ORGANISM: Human | | | | | | | | | |
| US-09-949-016-11312 | | | | | | | | | |
| Query Match | | | | | | | | | |
| Best Local Similarity 100.0%; Pred. No. 2.4e-181; | | | | | | | | | |
| Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | | | | | | |
| QY | 1 | MEMYQVKRYHGGAGLRLVELPTCMYRLPNVHGRSYGPAGAGHVEESNLISQALESRD | 60 | | | | | | |
| DB | 22 | MEMYQVKRYHGGAGLRLVELPTCMYRLPNVHGRSYGPAGAGHVEESNLISQALESRD | 81 | | | | | | |
| QY | 61 | DILKRLYLKKAAYDLSKICITPPADLDVTNIIQADEPTTLTNLDLSVSKYGA | 120 | | | | | | |
| DB | 82 | DILKRLYLKKAAYDLSKICITPPADLDVTNIIQADEPTTLTNLDLSVSKYGA | 141 | | | | | | |
| QY | 121 | DIVIANAPASPLSLIVLHRLICEHFRVLSVTHHTSSVKSVEENLLKCGEQNKOPROD | 180 | | | | | | |
| DB | 142 | DIVIANAPASPLSLIVLHRLICEHFRVLSVTHHTSSVKSVEENLLKCGEQNKOPROD | 201 | | | | | | |
| QY | 181 | YOLGFTLIWKVVPKTKQMKFSIQTMCPBIGEGNIARFLSLFCQKNAVNATLIDSWDIA | 240 | | | | | | |
| DB | 202 | YOLGFTLIWKVVPKTKQMKFSIQTMCPBIGEGNIARFLSLFCQKNAVNATLIDSWDIA | 261 | | | | | | |
| QY | 241 | IFOLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVLMVSYQOIGGCVTVPANVQ | 300 | | | | | | |
| DB | 262 | IFOLKEGSSKEKAAVFRSMNSALGKSPWLAGNELTVADVLMVSYQOIGGCVTVPANVQ | 321 | | | | | | |
| QY | 301 | RWRSCENLAPFNTALKLK 320 | | | | | | | |

Db 322 RMWRSCEMLAPFNTALKLK 341

RESULT 2

US-08-518-862C-2
Sequence 2, Application US/08518662C
Patent No. 5843757
GENERAL INFORMATION:
APPLICANT: Vogelstein, Bert
APPLICANT: Kinzler, Kenneth W.
APPLICANT: Nicolaides, Nicholas C.
TITLE OF INVENTION: Human JTV1 Gene Overlaps PMS2 Gene
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Ltd.
STREET: 1001 G Street, N.W.
CITY: Washington, D.C.
COUNTRY: U.S.A.
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/518, 862C
FILING DATE: 24-AUG-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107, 49697
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 312 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-518-862C-2

Query Match 97.8%; Score 1630; DB 2; Length 312;
Best Local Similarity 100.0%; Pred. No. 3.5e-177;
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPMYQKPYHGGGAPLRVLEPTCMYRLPNVHGRSYGAPAGAGHVOEESNLSIQLESROD 60
DB 1 MPMYQKPYHGGGAPLRVLEPTCMYRLPNVHGRSYGAPAGAGHVOEESNLSIQLESROD 60
QY 61 DILKRLYELKAAVNDGSKMIOPTPDADLVNTIIOADEPTTLTNNALDINSVYLGQYALK 120
DB 61 DILKRLYELKAAVNDGSKMIOPTPDADLVNTIIOADEPTTLTNNALDINSVYLGQYALK 120
QY 121 DIVIANNPSPPLSLVLRHLICSEFRVLSVTHSSVSKVSEPMILKCFGRONKKQPRQD 180
DB 121 DIVIANNPSPPLSLVLRHLICSEFRVLSVTHSSVSKVSEPMILKCFGRONKKQPRQD 180
QY 181 YOLGFTLLMKVNPCKTOMKFSIQTCPIEGEGNINARPLPSLFGQKHNAVNATLIDSWDIA 240
DB 181 YOLGFTLLMKVNPCKTOMKFSIQTCPIEGEGNINARPLPSLFGQKHNAVNATLIDSWDIA 240
QY 241 IFOLKEGSSKKAAPFRSMNSALGKSPMLAGNELTVADVLMVLSVLOOIGGCSVTIVANNO 300
DB 241 IFOLKEGSSKKAAPFRSMNSALGKSPMLAGNELTVADVLMVLSVLOOIGGCSVTIVANNO 300
QY 301 RMWRSCEMLAPF 312
DB 301 RMWRSCEMLAPF 312

RESULT 3

US-09-443-184-48
Sequence 48, Application US/09443184A
Patent No. 6372431
GENERAL INFORMATION:
APPLICANT: Cunningham, Mary Jane
APPLICANT: Zweiger, Gary
APPLICANT: Kaser, Matthew R.
APPLICANT: Panzer, Scott
APPLICANT: Selhammer, Jeffrey J.
APPLICANT: Yue, Henry
APPLICANT: Baughn, Mariah
APPLICANT: Azimzai, Yalda
APPLICANT: Lal, Preeti
TITLE OF INVENTION: MAMMALIAN TOXICOLOGICAL RESPONSE MARKERS
FILE REFERENCE: PC-0007 US
CURRENT FILING DATE: 1999-11-19
NUMBER OF SEQ ID NOS: 138
SOFTWARE: PERL Program
SEQ ID NO 48
LENGTH: 1512
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6372431 2302721CD1
US-09-443-184-48

Query Match 6.2%; Score 103; DB 3; Length 1512;
Best Local Similarity 22.1%; Pred. No. 0.12;
Matches 45; Conservative 30; Mismatches 69; Indels 60; Gaps 9;

QY 124 INANPASPPLSLVLRHLICSEFRVLSVTHSSVSKVSEPMILKCFGRONKKQPRQYQL 183
DB 6 LTVNSGDPPLGAL---LAVEHKQDVSI---SVBEGKENILH----- 41
QY 184 GFTLLMKVNPCKTOMKFSIQTCPIEGEGNINARPLPSLFGQKHNAVNATLIDSWV 237
DB 42 ---VSENVIFTDV-----NSILKRLARVATTGVLGS--NMHEHTEIDHWL 82
QY 238 DIAIFOLKEGSSKKAAPFRSMNSALGKSPMLAGNELTVADVLMVLSVLOOIGGCS----- 292
DB 83 EFSATKL--SSCDSFTSTINEINHLSTRTYLVGNSLSLADLCWATLK--GNAAMQEO 138
QY 293 ---VTVANVQRMWRSCEMLAPFNT 314
DB 139 KOKKAPVHVKEWFGFLAQQAFOG 162

RESULT 4

US-08-756-317-10
Sequence 10, Application US/08756317
Patent No. 5849894
GENERAL INFORMATION:
APPLICANT: Clemente, Thomas E.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Miskay, Timothy A.
APPLICANT: Stark, David M.
TITLE OF INVENTION: Improved Rhodospirillum Rubrum
TITLE OF INVENTION: Poly-B-Hydroxyalkanoate Synthase
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:57:26 ; Search time 50.2443 Seconds
(without alignments)
2084.158 Million cell updates/sec

Title: US-10-622-817-2

Perfect score: 1667

Sequence: 1 MPWVQKRYHGAGAPLRLVEL.....RWRNSCENTLAPPTALKLK 320

Scoring table:

BLASTSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBSCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBSCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBSCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBSCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBSCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBSCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBSCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBSCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBSCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBSCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBSCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBSCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBSCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------------------------|--------------------|
| 1 | 1630 | 97.8 | 312 | 16 US-10-463-676-4 | Sequence 4, Appli |
| 2 | 1191 | 71.4 | 229 | 16 US-10-463-676-6 | Sequence 5, Appli |
| 3 | 832 | 49.9 | 161 | 16 US-10-463-676-5 | Sequence 5, Appli |
| 4 | 102 | 6.1 | 201 | 15 US-10-369-493-13717 | Sequence 13717, A |
| 5 | 101 | 6.1 | 826 | 10 US-09-746-660A-92 | Sequence 92, Appli |
| 6 | 101 | 6.1 | 833 | 10 US-09-746-660A-90 | Sequence 90, Appli |
| 7 | 101 | 6.1 | 1221 | 9 US-09-919-891-2 | Sequence 2, Appli |
| 8 | 101 | 6.1 | 1221 | 9 US-09-919-891-5163 | Sequence 5163, Ap |
| 9 | 101 | 6.1 | 1221 | 15 US-10-450-055-2 | Sequence 2, Appli |
| 10 | 100.5 | 6.1 | 1221 | 17 US-10-494-675-156 | Sequence 156, App |
| 11 | 100.5 | 6.0 | 687 | 15 US-10-425-114-58897 | Sequence 58897, A |
| 12 | 100.5 | 6.0 | 1200 | 15 US-10-283-122A-53993 | Sequence 53993, A |
| 13 | 97.5 | 5.8 | 652 | 15 US-10-264-049-3001 | Sequence 3001, Ap |

| 14 | 97.5 | 5.8 | 711 | 16 | US-10-437-963-175483 | Sequence 175483, |
|----|------|-----|------|----|----------------------|--------------------|
| 15 | 97 | 5.8 | 559 | 9 | US-09-364-847-21 | Sequence 21, Appli |
| 16 | 97 | 5.8 | 856 | 9 | US-09-364-847-33 | Sequence 33, Appli |
| 17 | 97 | 5.8 | 856 | 9 | US-09-364-847-35 | Sequence 35, Appli |
| 18 | 97 | 5.8 | 5215 | 9 | US-09-861-289-2 | Sequence 2, Appli |
| 19 | 97 | 5.8 | 5215 | 9 | US-09-860-846-2 | Sequence 2, Appli |
| 20 | 97 | 5.8 | 5215 | 10 | US-09-988-348B-2 | Sequence 2, Appli |
| 21 | 97 | 5.8 | 5215 | 10 | US-09-836-821-2 | Sequence 45, Appli |
| 22 | 97 | 5.8 | 5215 | 14 | US-10-271-889-45 | Sequence 153870, |
| 23 | 96.5 | 5.8 | 914 | 16 | US-10-437-963-153870 | Sequence 29, Appli |
| 24 | 96.5 | 5.8 | 1148 | 16 | US-10-655-799-29 | Sequence 115189, |
| 25 | 96 | 5.8 | 606 | 16 | US-10-437-963-115189 | Sequence 42533, A |
| 26 | 95.5 | 5.7 | 455 | 15 | US-10-425-114-42633 | Sequence 1, Appli |
| 27 | 95 | 5.7 | 559 | 9 | US-09-821-016-1 | Sequence 1, Appli |
| 28 | 95 | 5.7 | 559 | 9 | US-09-820-952A-1 | Sequence 1, Appli |
| 29 | 95 | 5.7 | 559 | 14 | US-10-218-519-1 | Sequence 1, Appli |
| 30 | 95 | 5.7 | 559 | 14 | US-10-259-632-1 | Sequence 1, Appli |
| 31 | 95 | 5.7 | 559 | 14 | US-10-266-787-1 | Sequence 1, Appli |
| 32 | 95 | 5.7 | 559 | 14 | US-10-252-518-1 | Sequence 1, Appli |
| 33 | 95 | 5.7 | 559 | 17 | US-10-914-244-1 | Sequence 1, Appli |
| 34 | 95 | 5.7 | 722 | 15 | US-10-369-493-10013 | Sequence 10013, A |
| 35 | 95 | 5.7 | 936 | 14 | US-10-032-585-7442 | Sequence 7442, Ap |
| 36 | 94.5 | 5.7 | 955 | 15 | US-10-282-122A-69912 | Sequence 69912, A |
| 37 | 94 | 5.6 | 203 | 15 | US-10-369-493-8604 | Sequence 8604, Ap |
| 38 | 94 | 5.6 | 468 | 15 | US-10-424-599-260388 | Sequence 260388, |
| 39 | 94 | 5.6 | 1440 | 16 | US-10-437-963-132904 | Sequence 123904, |
| 40 | 93.5 | 5.6 | 569 | 16 | US-10-424-589-243766 | Sequence 243766, |
| 41 | 93.5 | 5.6 | 733 | 15 | US-10-424-599-251522 | Sequence 251522, |
| 42 | 92 | 5.5 | 739 | 15 | US-10-424-599-251522 | Sequence 63638, A |
| 43 | 92 | 5.5 | 1206 | 15 | US-10-283-122A-63838 | Sequence 30, Appli |
| 44 | 91.5 | 5.5 | 1149 | 16 | US-10-655-799-30 | Sequence 64644, A |
| 45 | 91.5 | 5.5 | 1192 | 15 | US-10-283-122A-64644 | |

ALIGNMENTS

RESULT 1
US-10-463-676-4
; Sequence 4, Application US/10463676
; Publication No. US20040175375A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Sunghoon
; TITLE OF INVENTION: Method for Treating Cancer Using P38/JTV-1 and Method
; FILE REFERENCE: 012679-09
; CURRENT APPLICATION NUMBER: US/10/463,676
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: KR 10-2003-13058
; PRIOR FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)...(312)
; OTHER INFORMATION: 1-312 amino acid sequence of p38/JTV-1
US-10-463-676-4
Query Match 97.8%; Score 1630, DB 16; Length 312;
Best Local Similarity 100.0%; Pred. No. 6.4e-153;
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPWVQKRYHGAGAPLRLVELPTQWRIPVHRSYGPAGAGHVEESLSIQALESSRD 60
DB 1 MPWVQKRYHGAGAPLRLVELPTQWRIPVHRSYGPAGAGHVEESLSIQALESSRD 60
QY 61 DIKLRLYEIKAVDGLSKMIGTPDADLVNTIIOADEPTTLTNALDLSVGLKQYGALK 120

Db 61 DILKRLYLKAAVDSGLSMIQTDPDADLVNTIIQADEPTTLTNALDINSVLGKDYALK 120
Qy 121 DIVINANPASPPLSLVLRLLCEHFRVLTSTVTHSSVSVPENLKCGEONKKQPRD 180
Db 121 DIVINANPASPPLSLVLRLLCEHFRVLTSTVTHSSVSVPENLKCGEONKKQPRD 180
Qy 181 YOLGFTLWKVNPKTOMKFSIQTMCPIDEGGNARLFLSIFGQKNAVNATLIDSWDIA 240
Db 181 YOLGFTLWKVNPKTOMKFSIQTMCPIDEGGNARLFLSIFGQKNAVNATLIDSWDIA 240
Qy 241 IFOLKEGSSKEKAAVFRSMNSALGKSPMLAGNELTVADVLSVLQOIGCSVTVPANVO 300
Db 241 IFOLKEGSSKEKAAVFRSMNSALGKSPMLAGNELTVADVLSVLQOIGCSVTVPANVO 300
Qy 301 RMRSCENLAPF 312
Db 301 RMRSCENLAPF 312

RESULT 2

US-10-463-676-6
; Sequence 6, Application US/10463676
; Publication No. US20040175375A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Sunghoon
; APPLICANT: Park, Bum-Joon
; TITLE OF INVENTION: Method for Treating Cancer Using P38/JTV-1 and Method
; FILE REFERENCE: 012679-091
; CURRENT APPLICATION NUMBER: US/10/463,676
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: KR 10-2003-13058
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)...(229)
; OTHER INFORMATION: 84-312 amino acid sequence p38/JTV-1
US-10-463-676-6

Query Match 71.4%; Score 1191; DB 16; Length 229;
Best Local Similarity 100.0%; Pred. No. 1.5e-109;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 DADLVNTIIQADEPTTLTNALDINSVLGKDYALKDIVINANPASPPLSLVLRLLC 143
Db 1 DADLVNTIIQADEPTTLTNALDINSVLGKDYALKDIVINANPASPPLSLVLRLLC 60
Qy 144 EHFRLVSTVHTSSVSVPENLKCGEONKKQPRDYOGLTLLMKVNPKTOMKFSIQ 203
Db 61 EHFRLVSTVHTSSVSVPENLKCGEONKKQPRDYOGLTLLMKVNPKTOMKFSIQ 120
Qy 204 MCPIDEGGNARLFLSIFGQKNAVNATLIDSWDIAIFOLKEGSSKEKAAVFRSMNSAL 263
Db 121 MCPIDEGGNARLFLSIFGQKNAVNATLIDSWDIAIFOLKEGSSKEKAAVFRSMNSAL 180
Qy 264 GKSPLMAGNELTVADVLSVLQOIGCSVTVPANVORMRSCENLAPF 312
Db 181 GKSPLMAGNELTVADVLSVLQOIGCSVTVPANVORMRSCENLAPF 229

RESULT 3

US-10-463-676-5
; Sequence 5, Application US/10463676
; Publication No. US20040175375A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Sunghoon
; APPLICANT: Park, Bum-Joon

; TITLE OF INVENTION: Method for Treating Cancer Using P38/JTV-1 and Method
; FILE REFERENCE: 012679-091
; CURRENT APPLICATION NUMBER: US/10/463,676
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: KR 10-2003-13058
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)...(161)
; OTHER INFORMATION: 1-161 amino acid sequence of p38/JTV-1
US-10-463-676-5

Query Match 49.9%; Score 832; DB 16; Length 161;
Best Local Similarity 100.0%; Pred. No. 3.9e-74;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPMYQVKEPHGGAFLRVELPTCMYRLPNVHGRSYPAPGAGHYQESNLSLOALESROD 60
Db 1 MPMYQVKEPHGGAFLRVELPTCMYRLPNVHGRSYPAPGAGHYQESNLSLOALESROD 60
Qy 61 DILKRLYLKAAVDSGLSMIQTDPDADLVNTIIQADEPTTLTNALDINSVLGKDYALK 120
Db 61 DILKRLYLKAAVDSGLSMIQTDPDADLVNTIIQADEPTTLTNALDINSVLGKDYALK 120
Qy 121 DIVINANPASPPLSLVLRLLCEHFRVLTSTVTHSSVSVK 161
Db 121 DIVINANPASPPLSLVLRLLCEHFRVLTSTVTHSSVSVK 161

RESULT 4

US-10-369-493-13717
; Sequence 13717, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 2002-02-21
; SEQ ID NO 13717
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Pseudomonas fluorescens
US-10-369-493-13717

Query Match 6.1%; Score 102; DB 15; Length 201;
Best Local Similarity 30.5%; Pred. No. 0.17;
Matches 32; Conservative 15; Mismatches 36; Indels 22; Gaps 3;

Qy 211 GNIRFLPSLFGQKNAVNA-TLIDSWDIAIFOLKEGSSKEKAAVFRSMNSALGKSPWL 269
Db 106 GPAARLITVGAFFNAEEVITRAHNFVKV-----MDLEFGKTPYL 146
Qy 270 AGNELTVADVLSVLQOIGCSVTVPANVORMRSCENLAPF 312
Db 147 AGTEPTIADVSAVSIAHAPGNVSLDDYANVRAMLRIEALPGF 191

A:Gene: ATSP:F2IE10.12
A:Map position: 5
A:Introns: 47/2; 141/1; 503/3; 659/3
C:Superfamily: yeast glutamate-tRNA ligase; glutamine-tRNA ligase homology
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis
F:223-499/Domain: glutamine-tRNA ligase homology <EGL>

Query Match 6.7%; Score 111; DB 2; Length 728;
Best Local Similarity 25.9%; Pred. No. 0.3;
Matches 48; Conservative 26; Mismatches 53; Indels 58; Gaps 10;

QY 128 PASPPLSLVLRLLCEHFRLVTHSVKSVPEMLKCGEQNKQPRODYOLGFTL 187
DB 10 PESPLSVYVALSLISASPV-----TIDSSAATVPSFV---FSDGRKLN-----GATV 55
QY 188 IMKNVPTQMKFSIQTMPIEGEINARFLSLFGQKNANVAT-----LIDSWD 238
DB 56 LIRIV-----GRSAKLPDPYG--NNAFDSQVSIILCINMKIDEMVD 95
QY 239 IAFOLKESKKAAVFRSMNSALGKSPWLAGNELTVADVLMASVLAQIGGCSVTVPA 297
DB 96 YASVF--SSGSEFENAC--GRVDKYLESSSTFLVGHSLSTADVALMSALAGTG----- 143
QY 298 NVQRW 302
DB 144 --QRW 146

RESULT 3
G82441
Probable glutathione S-transferase VCA0584 [imported] - vibrio cholerae (strain N16961)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: G82441
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gelin, M.L.; Dodson, R.J.; Chaudson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: G82441
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-222 <HRI>
A:Cross-references: UNIPROT:Q9RM05; GB:AB004389; GB:AB003853; NID:g9657989; PIDD:AAF9648
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Gene: VCA0584
A:Map position: 2
C:Superfamily: hypothetical protein b2302

Query Match 6.3%; Score 105.5; DB 2; Length 222;
Best Local Similarity 25.7%; Pred. No. 0.17;
Matches 28; Conservative 22; Mismatches 38; Indels 21; Gaps 4;

QY 197 MKFSIQTMPIEGEINARFLSLFGQKNANVATLIDSWDIAFLQKESKKAAY 256
DB 97 LRFQWGVPMWQANV---FTRYPEKIQPA---IDRY-----QKGRRLF 137
QY 257 RSMNSALGKSPWLAGNELTVADVLM--MSVLQIGGCSVTVPANVQRMW 303
DB 138 EVMDDQLAQNPLYAGDEYTIADATPPVRIHEMGSISIDGLTHLQRMK 186

RESULT 4
A29036
glutathione transferase (EC 2.5.1.18) Yb3 - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C:Accession: A29036
R:Abrahamovitz, M.; Listowsky, I.
J. Biol. Chem. 262, 7770-7773, 1987
A:Title: Selective expression of a unique glutathione S-transferase Yb3 gene in rat brain

A:Reference number: A29036; MUID:87222405; PMID:3584141
A:Accession: A29036
A:Molecule type: mRNA
A:Residues: 1-218 <ABR>
A:Cross-references: UNIPROT:P08009; GB:J02744; NID:g204512; PIDD:AAA41292.1; PID:g20451
C:Keywords: transferase

Query Match 5.8%; Score 97.5; DB 2; Length 218;
Best Local Similarity 25.6%; Pred. No. 0.78;
Matches 33; Conservative 24; Mismatches 41; Indels 31; Gaps 5;

QY 179 QDYQGFLLIMKNVPTQMKFSIQTMCPREGGINARFLSLFGQKNANVAT----- 231
DB 49 EKFKLG--LDFPNLP-----YLIDSHKITQSNALIRYL---GRKNHLCGTEBERIRV 97
QY 232 -----LIDSWDIA-----IFOLKESKKAAVFRSMNSALGKSPWLAGNELTVAD 278
DB 98 DLENGMDNMRVVALRLCYNPDPELTKPGYLEQLPGMKRLYSEFLGKAPWAGDKITVD 157
QY 279 VTLMSVLAQ 287
DB 158 FIAYDVLER 166

RESULT 5
JC5517
Gu/RNA helicase II binding protein - human
C:Species: Homo sapiens (man)
C>Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 05-Nov-1999
C:Accession: JC5517
R:Valdez, B.C.; Henning, D.; Perlsky, L.; Busch, R.K.; Busch, H.
Biochem. Biophys. Res. Commun. 234, 335-340, 1997
A:Title: Cloning and characterization of Gu/RH-II binding protein.
A:Reference number: JC5517; MUID:97320420; PMID:9177271
A:Accession: JC5517
A:Molecule type: mRNA
A:Residues: 1-645 <VAL>
A:Cross-references: GB:U78524; NID:g1696006; PIDD:AA58488.1; PID:g1696007
A:Note: It is uncertain whether Met-4 or Met-6 is the initiator
C:Comment: This protein is localized to the nucleus and interacts with Gu/RNA helicase I
C:Keywords: phosphoprotein
F:50-58/Region: nuclear location signal
F:362-374/Region: nuclear location signal
F:514-517,551-554,592-595,606-609/Region: 4-residue repeats (N-T-S-L)
F:7,177,322,412,432,460,461,462,470,614,627/Binding site: phosphate (Ser) (covalent) #sc
F:65,394,624/Binding site: phosphate (Thr) (covalent) #scatue predicted

Query Match 5.8%; Score 97.5; DB 2; Length 645;
Best Local Similarity 20.5%; Pred. No. 3.7;
Matches 79; Conservative 56; Mismatches 129; Indels 121; Gaps 20;

QY 16 LRV-ELPTCM-YRLPNVHGRSYGPARGAVQESNLSQALSRDDILKRLYEKAAY 73
DB 8 LRVSELQVLLGVAGRNHGRKHELLTYALHLRAG--CSPAVQMKIKEYRRRPPK-- 62
QY 74 DGLSKMIQTPDADLDVNIIOADEPTLTITNALDNLNSVKGQVAGLKQVYINNPASPP 133
DB 63 -----IMTP--ADLSTPNVHSSPMPTLSPS-----TIPQULTYDGHPRASSSE 102
QY 134 ---SL- VLHRLICEHF-RVLSVTHHSSVSVV-----ENILK--CFEQNKQ----- 176
DB 103 LPVSLGPKHELEPHLTSALHPVHVDIKQLKLPFDLIDELIKPTSLASDNGRFRERC 162
QY 177 -----PROYQGFLLIMKNVPTQMKFSIQ-----TMCPTEGGG----- 211
DB 163 PARALTPQVQVQSSSW--DISGTCDFTVQVQLRFLCSETSCPEQDHPPLCVKQNT 219
QY 212 -----NIARFLSLFGQKNANVATLIDSWV-----DI 239
DB 220 KPCSLPGYLPPTKGVGPKRPSRPINITSLV-----RLSTTVPTNTIVSWTAEIGNVSM 274
QY 240 AIRQLKEGSS-----KKAIVFRSMNSALGKSPWLAGNELTVADVLM--MSVLQIGGC 291

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:33:04 ; Search time 60.8145 Seconds
(without alignments)
2694.512 Million cell updates/sec

Title: US-10-622-817-2

Sequence: 1 MEMYGVKPRHGGAPLRLVEL.....RWRSCENLAPFNTALKLK 320

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|--------------|---------------------|
| 1 | 1667 | 100.0 | 320 | 1 MCA2_HUMAN | Q13155 homo sapien |
| 2 | 1663 | 99.8 | 320 | 2 O96CZ5 | O96CZ5 homo sapien |
| 3 | 1464 | 87.8 | 320 | 2 O8R010 | O8R010 mus musculu |
| 4 | 1464 | 87.8 | 320 | 2 O8R3Y6 | O8R3Y6 mus musculu |
| 5 | 1444 | 86.6 | 320 | 1 MCA2_CRIGR | Q9WVW7 cricetulus |
| 6 | 1260 | 75.6 | 280 | 2 O8R3V2 | O8R3V2 mus musculu |
| 7 | 1144.5 | 68.7 | 311 | 2 O6DK86 | O6DK86 xenopus tro |
| 8 | 1139.5 | 68.4 | 311 | 2 O6IND4 | O6IND4 xenopus lae |
| 9 | 1132.5 | 67.9 | 311 | 2 O7ZYD7 | O7ZYD7 xenopus lae |
| 10 | 887.5 | 53.2 | 321 | 2 O7Y3C0 | O7Y3C0 brachydanio |
| 11 | 267.5 | 16.0 | 340 | 2 O7Q7A3 | O7Q7A3 anopheles g |
| 12 | 224 | 13.4 | 301 | 2 O7KUM5 | O7KUM5 drosophila |
| 13 | 224 | 13.4 | 313 | 2 O8T060 | O8T060 drosophila |
| 14 | 224 | 13.4 | 322 | 2 O6NKM4 | O6NKM4 drosophila |
| 15 | 214.5 | 12.9 | 334 | 1 MCA2_DROME | O9VUR3 drosophila |
| 16 | 125.5 | 7.5 | 719 | 2 O82462 | O82462 arabidopsis |
| 17 | 111 | 6.7 | 728 | 2 O65253 | O65253 arabidopsis |
| 18 | 106.5 | 6.4 | 880 | 2 O8ERT1 | O8ERT1 oceanobacill |
| 19 | 106.5 | 6.4 | 913 | 2 O6DRB3 | O6DRB3 brachydanio |
| 20 | 106.5 | 6.4 | 913 | 2 O6PRQ4 | O6PRQ4 brachydanio |
| 21 | 105.5 | 6.3 | 222 | 2 O9KM05 | O9KM05 vibrio chol |
| 22 | 105 | 6.3 | 240 | 2 O8BTR1 | O8BTR1 mus musculu |
| 23 | 105 | 6.3 | 1512 | 1 SYEP_MOUSE | O8C9C7 mus musculu |
| 24 | 105 | 6.3 | 1800 | 2 O7PPA2 | O7PPA2 anopheles g |
| 25 | 104.5 | 6.3 | 922 | 2 O6T6M5 | O6T6M5 brachydanio |
| 26 | 104 | 6.2 | 661 | 2 O9VB85 | O9VB85 drosophila |
| 27 | 104 | 6.2 | 702 | 2 O8IRC6 | O8IRC6 drosophila |
| 28 | 104 | 6.2 | 841 | 2 O8IMQ2 | O8IMQ2 drosophila |
| 29 | 103 | 6.2 | 328 | 2 O86X73 | O86X73 homo sapien |
| 30 | 103 | 6.2 | 869 | 2 O6PD57 | O6PD57 homo sapien |
| 31 | 103 | 6.2 | 1213 | 2 O8FTD2 | O8FTD2 corynebacte |

| | | | | | |
|----|-------|-----|------|--------------|--------------------|
| 32 | 102.5 | 6.1 | 372 | 2 O87XN7 | O87XN7 pseudomonas |
| 33 | 102.5 | 6.1 | 715 | 2 O9LIZ8 | O9LIZ8 oryza sativ |
| 34 | 102 | 6.1 | 559 | 2 O8KQ23 | O8KQ23 pseudomonas |
| 35 | 102 | 6.1 | 559 | 2 O88D25 | O88D25 pseudomonas |
| 36 | 102 | 6.1 | 1486 | 2 O6TXE9 | O6TXE9 ratius norv |
| 37 | 101 | 6.1 | 1196 | 2 O6M580 | O6M580 corynebacte |
| 38 | 101 | 6.1 | 1221 | 2 O8NQD1 | O8NQD1 corynebacte |
| 39 | 100.5 | 6.0 | 454 | 2 O6UDR4 | O6UDR4 mus musculu |
| 40 | 100.5 | 6.0 | 1201 | 2 O6NH83 | O6NH83 corynebacte |
| 41 | 99.5 | 6.0 | 651 | 1 P1A1_MOUSE | O88907 mus musculu |
| 42 | 99 | 5.9 | 203 | 2 O6W112 | O6W112 rhizobium b |
| 43 | 98.5 | 5.9 | 222 | 2 O8EGM4 | O8EGM4 shewanella |
| 44 | 98.5 | 5.9 | 582 | 2 O66CB1 | O66CB1 yersinia ps |
| 45 | 97.5 | 5.8 | 217 | 1 GTW3_RAT | P08009 ratius norv |

ALIGNMENTS

RESULT 1
MCA2_HUMAN STANDARD; PRT; 320 AA.
ID Q13155; O9P1L2;
AC Q13155; O9P1L2;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Multisynthetase complex auxiliary component p38 (JTV-1 protein)
DE (PRO0992).
GN Name=JTV1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96115582; PubMed=8666379;
RA Nicolaides N.C., Kinzler K.W., Vogelstein B.;
RT "Analysis of the 5' region of PMS2 reveals heterogeneous transcripts
and a novel overlapping gene.";
RL Genomics 29:329-334(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Lymph;
MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.T., Utidin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.E., Garcia A.M., Gay L.J., Hultk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
FAhey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE OF 197-320 FROM N.A.
RC TISSUE=Fetal liver;
Zhang C., Yu Y., Zhang S., Wei H., Zhou G., Ouyang S., Luo L., Bi J.,
Liu M., He F.;
RT "Functional prediction of the coding sequences of 121 new genes
deduced by analysis of cDNA clones from human fetal liver.";
Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
[4]

RP INTERACTION WITH FUBP1.
RX MEDLINE=22716800; PubMed=12819782; DOI=10.1038/ng1182;
RA Kim M.J., Park B.-J., Kang Y.-S., Kim H.J., Park J.-H., Kang J.W.,
RT Lee S.W., Han J.M., Lee H.-W., Kim S.;
RT "Downregulation of FUBP-binding protein and c-myc by tRNA synthetase
RT cofactor p38 is required for lung cell differentiation.";
RL Nat. Genet. 34:330-336(2003).
CC -1- FUNCTION: Probable core protein of the multisynthetase complex
CC structure. Mediates ubiquitination of FUBP1 and its degradation by
CC the proteasome.
CC -1- SUBUNIT: Component of the multisynthetase complex which is
CC comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the
CC monospecific isoleucyl, leucyl, glutamyl, methionyl, lysyl,
CC arginyl, and aspartyl-tRNA synthetases as well as three auxiliary
CC proteins, p18, p48 and p43. Binds FUBP1.
CC -1- SIMILARITY: Contains 1 GST-like domain.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 312.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, U24169; AAC50391.1; ALT FRAME.
DR EMBL, BC002853; AAH02853.1; -
DR EMBL, BC010156; AAH01056.1; -
DR EMBL, AF116615; AAF71039.1; -
DR H-InvDB; HIX0006460; -
DR MIM; 600859; -
DR InterPro; IPR010987; GST_C like.
DR InterPro; IPR004046; GST_Cterm.
DR Pfam; PF00043; GST_C_1.
DR KX Protein biosynthesis.
SQ SEQUENCE 320 AA; 35349 MW; F253726B63C12BAB CRC64;
Query Match 100.0%; Score 1667; DB 1; Length 320;
Best Local Similarity 100.0%; Pred. No. 6e-127;
Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPWYQVRYHGGAGPLRVLELPTCMYRLPNVHGSRGYPAPAGHVOEESNLSLOALESROD 60
DB 1 MPWYQVRYHGGAGPLRVLELPTCMYRLPNVHGSRGYPAPAGHVOEESNLSLOALESROD 60
QY DILKRLVELKAAVDSLSKMIQTDPADLDVTNIIQADEPTTLTTNALDINSVLGKDYGALK 120
DB 61 DILKRLVELKAAVDSLSKMIQTDPADLDVTNIIQADEPTTLTTNALDINSVLGKDYGALK 120
QY 121 DIVINANPASPPLSLVLRHLCEHFRVLSVTHSSSVKSPENLLKCRGEQNKQPROD 180
DB 121 DIVINANPASPPLSLVLRHLCEHFRVLSVTHSSSVKSPENLLKCRGEQNKQPROD 180
QY 121 DIVINANPASPPLSLVLRHLCEHFRVLSVTHSSSVKSPENLLKCRGEQNKQPROD 180
DB 121 DIVINANPASPPLSLVLRHLCEHFRVLSVTHSSSVKSPENLLKCRGEQNKQPROD 180
QY 181 YOLGFTLLIMKNVPTQMKFSIQTMCPIDEGBNIAFLFSLFGQKNAVNAATLIDSWDIA 240
DB 181 YOLGFTLLIMKNVPTQMKFSIQTMCPIDEGBNIAFLFSLFGQKNAVNAATLIDSWDIA 240
QY 241 IFOLKEGSSKEKAAVFRSMNSALGKSPMLAGNELTVADVILMSVLOQIGGCSVTVPANVQ 300
DB 241 IFOLKEGSSKEKAAVFRSMNSALGKSPMLAGNELTVADVILMSVLOQIGGCSVTVPANVQ 300
QY 301 RWRKSCENLAPENTALKLK 320
DB 301 RWRKSCENLAPENTALKLK 320
RESULT 2
ID 096CZ5 PRELIMINARY; PRT; 320 AA.
AC 096CZ5;

DT 01-DEC-2001 (TRENBLREL 19, Created)
DT 01-DEC-2001 (TRENBLREL 19, Last sequence update)
DT 01-MAR-2004 (TRENBLREL 26, Last annotation update)
DE JTV1.
GN Name=JTV1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.U., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lomellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., Mewhan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz U., Myers R.M., Butlerfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL, BC013630; AAH13630.1; -
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR010987; GST_C like.
DR Pfam; PF00043; GST_C_1.
SQ SEQUENCE 320 AA; 35335 MW; 19F14BF756612E08 CRC64;
Query Match 99.8%; Score 1663; DB 2; Length 320;
Best Local Similarity 99.7%; Pred. No. 1.3e-126;
Matches 319; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MPWYQVRYHGGAGPLRVLELPTCMYRLPNVHGSRGYPAPAGHVOEESNLSLOALESROD 60
DB 1 MPWYQVRYHGGAGPLRVLELPTCMYRLPNVHGSRGYPAPAGHVOEESNLSLOALESROD 60
QY 61 DILKRLVELKAAVDSLSKMIQTDPADLDVTNIIQADEPTTLTTNALDINSVLGKDYGALK 120
DB 61 DILKRLVELKAAVDSLSKMIQTDPADLDVTNIIQADEPTTLTTNALDINSVLGKDYGALK 120
QY 121 DIVINANPASPPLSLVLRHLCEHFRVLSVTHSSSVKSPENLLKCRGEQNKQPROD 180
DB 121 DIVINANPASPPLSLVLRHLCEHFRVLSVTHSSSVKSPENLLKCRGEQNKQPROD 180
QY 121 DIVINANPASPPLSLVLRHLCEHFRVLSVTHSSSVKSPENLLKCRGEQNKQPROD 180
DB 121 DIVINANPASPPLSLVLRHLCEHFRVLSVTHSSSVKSPENLLKCRGEQNKQPROD 180
QY 181 YOLGFTLLIMKNVPTQMKFSIQTMCPIDEGBNIAFLFSLFGQKNAVNAATLIDSWDIA 240
DB 181 YOLGFTLLIMKNVPTQMKFSIQTMCPIDEGBNIAFLFSLFGQKNAVNAATLIDSWDIA 240
QY 241 IFOLKEGSSKEKAAVFRSMNSALGKSPMLAGNELTVADVILMSVLOQIGGCSVTVPANVQ 300
DB 241 IFOLKEGSSKEKAAVFRSMNSALGKSPMLAGNELTVADVILMSVLOQIGGCSVTVPANVQ 300
QY 301 RWRKSCENLAPENTALKLK 320
DB 301 RWRKSCENLAPENTALKLK 320

```
RESULT 3
OBR010 PRELIMINARY; PRT; 320 AA.
ID OBR010
AC OBR010;
DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE Jtvl-pending protein.
GN Name-Jtvl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N, and FVB/N-3; TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshlyuk S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McGwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richardson D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N-3; TISSUE=Mammary tumor;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024410; AAH24410.1; -.
DR EMBL; BC026972; AAH26972.1; -.
DR MGD; MGI:2385237; Jtvl.
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR010987; GST_C_1like.
DR Pfam; PF00043; GST_C_1.
SQ SEQUENCE 320 AA; 35396 MW; 1745D7EE4BC3670D CRC64;

Query Match 87.8%; Score 1464; DB 2; Length 320;
Best Local Similarity 87.2%; Pred. No. 1.8e-110;
Matches 279; Conservative 16; Mismatches 25; Indels 0; Gaps 0;
```

```
DB 181 YOLGFTLWKVNPKTOMKFSVQTMCPIEEGNIARFLFSLFGQKHNAVTLTIDSWDIA 240
QY 241 IFOLKSGSKERAAVRSNNSALGKSPMLAGNELTVADVLNLSVLQIGGCSVTVPANQ 300
DB 241 IFOLKSGSKERAAVRSNNSALGKSPMLAGNELTVADVLNLSVLQIGGCSVTVPANQ 300
QY 301 RMRKSCENLAPNTALKLK 320
DB 301 RMRKSCENLAPNTALKLK 320

RESULT 4
OBR2Y6 PRELIMINARY; PRT; 320 AA.
ID OBR2Y6
AC OBR2Y6;
DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Jtvl-pending protein.
GN Name-Jtvl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II; TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshlyuk S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McGwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richardson D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II; TISSUE=Mammary tumor;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC026958; AAH26958.1; -.
DR MGD; MGI:2385237; Jtvl.
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR010987; GST_C_1like.
DR Pfam; PF00043; GST_C_1.
SQ SEQUENCE 320 AA; 35423 MW; 1C21FLA74C9882B4 CRC64;

Query Match 87.8%; Score 1464; DB 2; Length 320;
Best Local Similarity 87.2%; Pred. No. 1.8e-110;
Matches 279; Conservative 16; Mismatches 25; Indels 0; Gaps 0;
```

```

Db      121 DIVINANPASPPLSLVLRLLCERFVSLVTHSSVKNVENVLVCFGEOARKOSRHE 180
Qy      181 YOLGFTLLMKVNPVKTKQKFSIQTMCPPIEGEGNIARFLFSLFGQKKNVANAATLIDSWDIA 240
Db      181 YOLGFTLLMKVNPVKTKQKFSVQTMCPPIEGEGNIARFLFSLFGQKKNVANAATLIDSWDIA 240
Qy      241 IFOLKEGSSKEKAIVFRSNNSALGKSPMLAGNELTVADVLMVSLVLOQTGGCSVTPVAVNQ 300
Db      241 MFOLKEGSSKEKAIVFRSNNSALGKSPMLVGNELTVADVLMVSLVLOQTGGCSGVAAPVAVNQ 300
Qy      301 RMWRCENLAPFPTALKLK 320
Db      301 RMWRCENLAPFPTALKLK 320

```

```

RESULT 5
MCA2_CRIGR STANDARD; PRT; 320 AA.
AC 09WVW7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Multisynthetase complex auxiliary component p38.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetus.
OX NCBI_TaxID=10029;
RN 1;
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=99096915; PubMed=9878398; DOI=10.1006/jmbi.1998.2316;
RA Quesvillon S., Robinson J.-C., Berthomieu E., Slatock M., Mirande M.;
RT "Macromolecular assemblage of aminoacyl-tRNA synthetases:
RT identification of protein-protein interactions and characterization of
RT a core protein."
RL J. Mol. Biol. 285:183-195(1999).

```

```

CC -1- FUNCTION: Probable core protein of the multisynthetase complex
CC that serves as a template for the assembly of the supramolecular
CC structure.
CC -1- SUBUNIT: Component of the multisynthetase complex which is
CC comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the
CC monospecific isoleucyl, leucyl, glutamyl, methionyl, lysyl,
CC aspartyl, and asparaginyl-tRNA synthetases as well as three auxiliary
CC proteins, p18, p48 and p43.
CC -1- SIMILARITY: Contains 1 GST-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF072727; AAD38422.1; -
DR InterPro: IPR010987; GST_C_1like.
DR InterPro: IPR004046; GST_Cterm.
DR Pfam: PF00043; GST_C; 1.
DR KMW Protein biosynthesis.
SQ SEQUENCE 320 AA; 35433 MW; 6D24E033ABEC10A CRC64;

```

```

Query Match      86.6%; Score 1444; DB 1; Length 320;
Best Local Similarity 86.6%; Pred. No. 7.6e-109;
Matches 277; Conservative 16; Mismatches 27; Indels 0; Gaps 0;

```

```

Qy      1 MEMYOVKPYHGAGLAVELPTCYRLPNVGRSVGPAPGAGRVQESNSISLOAESRQD 60
Db      1 MEMYOVKSYHGAGLAVELPTCYRLPNVGRSVKTSPTVAGVQESNSISLOAESRQD 60
Qy      61 DILKRLYLKAAVNGLSKMIOTPDADLVNIIQADEPTLTITNALDINSVGLKQYALK 120
Db      61 DILKRLYLKAAVNGLSKMIOTPDADLVNIIQADEPTLTITNALDINSVGLKQYALK 120

```

```

Db      61 DILKRLYLKAAVNGLSKMIOTPDADLVNIIQADEPTLTITNALDINSVGLKQYALK 120
Qy      121 DIVINANPASPPLSLVLRLLCERFVSLVTHSSVSNVENVLVCFGEOARKOSRHE 180
Db      121 DIVINANPASPPLSLVLRLLCERFVSLVTHSSVSNVENVLVCFGEOARKOSRHE 180
Qy      181 YOLGFTLLMKVNPVKTKQKFSIQTMCPPIEGEGNIARFLFSLFGQKKNVANAATLIDSWDIA 240
Db      181 YOLGFTLLMKVNPVKTKQKFSVQTMCPPIEGEGNIARFLFSLFGQKKNVANAATLIDSWDIA 240
Qy      241 IFOLKEGSSKEKAIVFRSNNSALGKSPMLAGNELTVADVLMVSLVLOQTGGCSVTPVAVNQ 300
Db      241 MFOLKEGSSKEKAIVFRSNNSALGKSPMLVGNELTVADVLMVSLVLOQTGGCSGVAAPVAVNQ 300
Qy      301 RMWRCENLAPFPTALKLK 320
Db      301 RMWRCENLAPFPTALKLK 320

```

```

RESULT 6
Q8R3V2 PRELIMINARY; PRT; 280 AA.
ID Q8R3V2;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE JTV1.
GN Name=Jtv1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butcherfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerker A., Schein J.E.,
RA Jones S.J., Merra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA Straubeberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC024480; AAH24480.1; -
DR MGD: WGI:2385237; JTV1.
DR InterPro: IPR004046; GST_Cterm.
DR InterPro: IPR010987; GST_C_1like.
DR Pfam: PF00043; GST_C; 1.
SQ SEQUENCE 280 AA; 31097 MW; E5BE81498983FD2E CRC64;

```

```

Query Match      75.6%; Score 1260; DB 2; Length 280;
Best Local Similarity 88.0%; Pred. No. 5.5e-94;
Matches 243; Conservative 14; Mismatches 19; Indels 0; Gaps 0;

```

```

Qy      45 QEESNLSIQALIESRQODILKRLYLKAAVNGLSKMIOTPDADLVNIIQADEPTLTITN 104

```

```
Db 5 QERSESLQALLESRODILKRLVELKAAVDGLSKMHTPDADLVNTIIQADEPTLATY 64
Qy 105 ALDLNVLGKDYALKDYVNNANPASPPLSLVLRLLCEHFRVLTWHTSSVAVPEN 164
Db 65 TLDLNVLGKDYALKDYVNNANPASPPLSLVLRLLCEHFRVLTWHTSSVAVPEN 124
Qy 165 LKCFPEBONKKOPRODYQAGFTLIMKNVPKTKMFKSIOTMCPTEGEGNIARPLFSLFGOK 224
Db 125 LVNCFPEBQARKSRHEYLQAGFTLIMKNVPKTKMFKSIOTMCPTEGEGNIARPLFSLFGOK 184
Qy 225 HNAVNATLIDSWVDIAIFOLKEGSSKEKAAVFRSMNSALGKSPMLAGNELTVADVLMVY 284
Db 185 HNAVNATLIDSWVDIAMFQREGSSKEKAAVFRSMNSALGSRPMVIGNELTVADVLMVY 244
Qy 285 LQOIGGCSVTVPANVORMRSCENLAPFTALKLLK 320
Db 245 LQOIGGSSGAAPTNVORMLKSCENLAPFTALQLLK 280
```

RESULT 7

```
Qy 06DK86 PRELIMINARY; PRT; 311 AA.
ID 06DK86
AC 25-OCT-2004 (TEMBLrel. 28, Created)
DT 25-OCT-2004 (TEMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TEMBLrel. 28, Last annotation update)
DE MGC69221 protein.
GN Name=MGC69221;
OC Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
NCBI_TaxID=8364;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Tohyuki S., Carninci P., Prange C.,
RA Raha S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywicki M.I., Skalska U., Smalms D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.;
RA Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC074561; AAH74561.1; -.
DR InterPro; IPR004046; GST_C_1like.
DR InterPro; IPR010987; GST_C_1like.
DR Pfam; PF00043; GST_C_1.
SQ SEQUENCE 311 AA; 34480 MW; D98F27F73C466154 CRC64;
```

Query Match 68.7%; Score 1144.5; DB 2; Length 311;
Best Local Similarity 68.8%; Pred. No. 1.5e-84;
Matches 220; Conservative 43; Mismatches 48; Indels 9; Gaps 3;

```
Qy 1 MBWQVQPYHGGAGPLRVELPTCMYRLPNVHGRSGYPAGAGHVOESNLQALLESROD 60
Db 1 MBWQVQPYCGG--ELQVLDPTCMYRLPNVH-----QAVSENLEGAADPAIQALESROG 52
Qy 61 DILKRLYELKAAVDGLSKMHTPDADLVNTIIQADEPTLTALDNLVSLGKDYGALK 120
Db 53 DILKRLYELKAAVDGLSKMHTPDADLVNTIIQADEPTLTALDNLVSLGKDYGALK 112
Qy 121 DIVYNNANPASPPLSLVLRLLCEHFRVLTWHTSSVAVPENLKCFCGEQKKOPROD 180
Db 113 DIVYNNANPASPPLSLVLRLLCEHFRVLTWHTSSVAVPENLKCFCGEQKKOPROD 172
Qy 181 YQAGFTLIMKNVPKTKMFKSIOTMCPTEGEGNIARPLFSLFGOKHNAVNATLIDSWVDIA 240
Db 173 YQAGFTLIMKNVPKTKMFKSIOTMCPTEGEGNIARPLFSLFGOKHNAVNATLIDSWVDIA 232
Qy 241 IFOLKEGSSKEKAAVFRSMNSALGKSPMLAGNELTVADVLMVYLQOIGGCSVTVPANVQ 300
Db 233 IFOLREGSSKEKAAVFRSMNSALGKSPMLAGNELTVADVLMVYCAVQCGN-STAVPSPVQ 291
Qy 301 RMRSCENLAPFTALKLLK 320
Db 292 KMRKSCENLAPFTALKLLK 311
```

RESULT 8

```
Qy 06IND4 PRELIMINARY; PRT; 311 AA.
ID 06IND4
AC 05-JUL-2004 (TEMBLrel. 27, Created)
DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE MGC80304 protein.
GN Name=MGC80304;
OC Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
NCBI_TaxID=8355;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Tohyuki S., Carninci P., Prange C.,
RA Raha S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywicki M.I., Skalska U., Smalms D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
Dev. Dyn. 225:384-391(2002).
[3]
```


RP SEQUENCE FROM N.A.

RC TISSUE=Ovary;

RA Klein S., Gerhard D.S.;

RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL: BC072178; AAH2178.1; -

DR InterPro: IPR004046; GST_Cterm.

DR InterPro: IPR010987; GST_C_1like.

DR Pfam: PF00043; GST_C_1.

DR S0 SEQUENCE 311 AA; 34424 MW; D5E8325C18D86751 CRC64;

Query Match

Best Local Similarity 68.4%; Score 1139.5; DB 2; Length 311;
Matches 221; Conservative 40; Mismatches 50; Indels 9; Gaps 3;

QY 1 MEMYQKPYHGGAPRLVELPTCMYRLPNVHSGSYGAPAGHVOESNLSLOAESROD 60

DB 1 MEMYQKPYHGGAPRLVELPTCMYRLPNVHSGSYGAPAGHVOESNLSLOAESROD 60

QY 61 DILKRLVELKAAVAGLSKMTQTPDADLVTHIIQADEPTTLTNNALDINSVIGKYGALK 120

DB 53 DILKRLVELKAAVAGLSKMTQTPDADLVTHIIQADEPTTLTNNALDINSVIGKYGALK 112

QY 121 DIVINANPASPPLSLVLRHLCEHFRVSTVHTSSVSVENILKCFGEONKKOPROD 180

DB 113 DIVINANPASPPLSLVLRHLCEHFRVSTVHTSSVSVENILKCFGEONKKOPROD 172

QY 181 YOLGFTLIWKNVKPTQMKFSIQTMCPTEEGNIAFLFSLFQKHNANATLIDSWDIA 240

DB 173 YOLGFTLIWKNVKPTQMKFSIQTMCPTEEGNIAFLFSLFQKHNANATLIDSWDIA 232

QY 241 IFOLKSGSSKEKAVERSNMNSALGKSPMLAGNELTVADVLMSVLOQIGGCSVTVPANVO 300

DB 233 IFOLKSGSSKEKAVERSNMNSALGKSPMLAGNELTVADVLMSVLOQIGGCSVTVPANVO 291

QY 301 RWRSGCENLAPPTALKLLK 320

DB 292 KMKKSCENLAPPTALKLLK 311

RESULT 9

Q72YD7 PRELIMINARY; PRT; 311 AA.

AC 072YD7; MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

DT 01-JUN-2003 (TREMBLrel. 24, Created)

DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)

DE 01-MAR-2004 (TREMBLrel. 26, Last annotation update)

OS Jtrv1-prov protein.

OC Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;

OC Xenopodinae; Xenopus.

NCBI_TaxID=8355;

RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Brownstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Villalón D.K., Muray D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Krzywinski M.I., Skalska U., Small D.E., Schermer A., Schein J.E.,

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;

RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,

RA Richardson P.,

RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus

RA Dev. Dyn. 225:384-391(2002).

Query Match

Best Local Similarity 67.9%; Score 1132.5; DB 2; Length 311;
Matches 221; Conservative 39; Mismatches 51; Indels 9; Gaps 4;

QY 1 MEMYQKPYHGGAPRLVELPTCMYRLPNVHSGSYGAPAGHVOESNLSLOAESROD 60

DB 1 MEMYQKPYHGGAPRLVELPTCMYRLPNVHSGSYGAPAGHVOESNLSLOAESROD 60

QY 61 DILKRLVELKAAVAGLSKMTQTPDADLVTHIIQADEPTTLTNNALDINSVIGKYGALK 120

DB 53 DILKRLVELKAAVAGLSKMTQTPDADLVTHIIQADEPTTLTNNALDINSVIGKYGALK 112

QY 121 DIVINANPASPPLSLVLRHLCEHFRVSTVHTSSVSVENILKCFGEONKKOPROD 180

DB 113 DIVINANPASPPLSLVLRHLCEHFRVSTVHTSSVSVENILKCFGEONKKOPROD 172

QY 181 YOLGFTLIWKNVKPTQMKFSIQTMCPTEEGNIAFLFSLFQKHNANATLIDSWDIA 240

DB 173 YOLGFTLIWKNVKPTQMKFSIQTMCPTEEGNIAFLFSLFQKHNANATLIDSWDIA 232

QY 241 IFOLKSGSSKEKAVERSNMNSALGKSPMLAGNELTVADVLMSVLOQIGGCSVTVPANVO 300

DB 233 IFOLKSGSSKEKAVERSNMNSALGKSPMLAGNELTVADVLMSVLOQIGGCSVTVPANVO 291

QY 301 RWRSGCENLAPPTALKLLK 320

DB 292 KMKKSCENLAPPTALKLLK 311

RESULT 10

Q72YD7 PRELIMINARY; PRT; 321 AA.

AC 072YD7; MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

DT 01-OCT-2003 (TREMBLrel. 25, Created)

DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)

DE 01-MAR-2004 (TREMBLrel. 26, Last annotation update)

OS ORFpames=296:63976;

OC Brachydanio rerio (Zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;

OC Cyprinidae; Danio.

NCBI_TaxID=7955;

RP SEQUENCE FROM N.A.

RC TISSUE=Kidney;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Brownstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Villalón D.K., Muray D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Krzywinski M.I., Skalska U., Small D.E., Schermer A., Schein J.E.,

RA Hopkings R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Honig L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Uedtin T.B., Toohily S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., Mcwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywicki M.I., Skalska U., Smallos D.E., Scherch B., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC053178; AAH53178.1; -
DR ZFIN: ZDB-GENE-040426-2652; zgc:63976.
DR InterPro: IPR004046; GST_Cterm.
DR InterPro: IPR010987; GST_C_1like.
DR Pfam: PF00043; GST_C_1; 34852 MW; BA8F6B951208244A CR64;
SQ SEQUENCE 321 AA; 34852 MW; BA8F6B951208244A CR64;
Query Match 53.2%; Score 887.5; DB 2; Length 321;
Best Local Similarity 55.6%; Pred. No. 1,1e-63;
Matches 184; Conservative 50; Mismatches 76; Indels 21; Gaps 8;
QY 1 MPMYQVKEPHGGAPLRLEPTCMYRLPNVHGSRYPAGAGVQF-ESNLSIQALESRQ 59
DB 1 MPMYQVKEPV--SPADTVLPLPTCMYRLPNVHAQ--GASLGEHLQVGEVDPYKALAEHQ 56
QY 60 DDLIKLYELKAVDGLSKMIOPTPDADLVNTNIIQADBPPTLL-----TNALDLSV 111
DB 57 DELIKLYELKAVDGLAKVTTPDADLDLSTLAH-----TLNHTADANVLRSTADLDL 111
QY 112 LGGDYGALDIVINANPASPPLSLVLRLLCEHFEVLSIVHTHSIVKSPENILKCFGS 171
DB 112 LGGDYGALDIVINANPAPPLSLVLRLLCGRFVLSVHNSIVSTPAPPLSCLSP 171
QY 172 QN-KKQPRDDYQLGFTLIMKNVPTQMKFSIQMCFIEBEGNIARLFSLFG-QKGNAYN 229
DB 172 RHTHSYARRRFPQGLFTLIMKDVSKLQMKFSTQMKPIEGGNVARRFLYRLGAEPRDPVS 231
QY 230 ATLIDSVNDIAIFQLKEGSSKEKAARFRSMNSALGKSPMLAGNELTVADVLMVSLVQIQG 289
DB 232 ATLMDGVNVTALFQLAEGGSKKAAVLRALNALGRSPMLLGGFSLDVSACVLAQIQG 291
QY 290 GCSVTVPANVQWRMRSCENLAPNTALKLK 320
DB 292 QTS-SAPANVQWRKLSCKQNLGYRSCVDPILQ 321
RESULT 11
Q707A3 PRELIMINARY; PRT; 340 AA.
AC Q707A3;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AGCP808 (Fragment).
GN Name=agc5014; ORFName=ENSG000000011827;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCB1_Taxid=180454;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AAB01008960; BAA10870.1; -
FT NON TER 1 1
SQ SEQUENCE 340 AA; 37467 MW; FDD0B525CD4E9976 CR64;
Query Match 16.0%; Score 267.5; DB 2; Length 340;
Best Local Similarity 29.2%; Pred. No. 2.6e-13;
Matches 100; Conservative 47; Mismatches 133; Indels 63; Gaps 13;
QY 3 MPMYQVKEPHGGAPLRLEPTCMYRLPNVHGSRYP-----APG---AGHV----- 44
DB 12 MYRLKPVW--NODVCELPCTMYTLTPVCAVSHDPAGFADSDTGTGLAAGVDPMSIEID 69
QY 45 ---QEEENLSIQALESRQDDILKRYELKAVDGLSKMIOPTPDADLV--TNIIQADEPTT 100
DB 70 NLLQKAEDELMLALRQQRVILQQLAEK-----KEIMARTELKLNANPAPVQPS 122
QY 101 -----LTTNALDLSVIGKDYGLKDIVINANPASPPLSLVLRLLCEHFRVLSYVA 153
DB 123 PLKSKAQLKAEPIINLT-----CLQDFVNVAPEYVYSILATLKNLWKDRNLNQEGF 174
QY 154 THSSVYSVE-----NLLKCFGEQNKQPRDDYQGLFTLIMKNV--PTQMKFSIQTMCP 206
DB 175 THSTVPEKSEBALAFQNAVTASGTAAANLPR-----IKVTLIMKNVGAIVEMTSPSYVP 230
QY 207 IEGEGNIARFLPSLFGQKIN---AVNATLIDSVNDIAIFQLKEGSSKEKAARFRSMNSAL 263
DB 231 ICGEVNIALRYLRCGSEFENYEQDNVDEVDLIDLCVYLKKNQKQKQILRTIGAKL 290
QY 264 GSPWLAGNELTVADVLMVSLVQIQGCGSVTPANVQWRMRSC 306
DB 291 GRAAGFGADLSLDCIAFTSAVKQV---QRTVAKDVNPMNHKC 330
RESULT 12
Q7KUM5 PRELIMINARY; PRT; 301 AA.
AC Q7KUM5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE CG12304-PB.
GN ORFName=CG12304;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCB1_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=201956006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazel R.G., Chame M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter B.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Berryklatoglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimball B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laeko P., Lei Y., Levisky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Mostreli A.,
 RA Mount S.M., Moy M., Murphy B., Murphy C., Morris J., Mostreli A.,
 RA Nelson D.R., Nelson K.A., Nixon K., Pollard J., Puri V., Reese M.G.,
 RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhu X., Yao Q.A., Ye J.,
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [2]

RP SEQUENCE FROM N.A.
 RX MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Change M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
 RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
 RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*
 RT *melanogaster* euchromatic genome sequence.";
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]

RP SEQUENCE FROM N.A.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
 RT a genomic perspective.";
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 RN [4]

RP SEQUENCE FROM N.A.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.V., Bayraktaroglu L., Berman B.P.,
 RA Betencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RN [5]

RP SEQUENCE FROM N.A.
 RG FlyBase;
 RN Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RG FlyBase;
 RN Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AEO05530; AAS64998.1; -;
 DR InterPro; IPR004046; GST_Cterm.
 DR InterPro; IPR010987; GST_C_1like.
 DR Pfam; PF00043; GST_C_1.
 SQ SEQUENCE 301 AA; 33241 MW; 35224E17FE03F3BE CRC64;

Query Match 13.4%; Score 224; DB 2; Length 301;
 Best Local Similarity 28.1%; Pred. No. 7.6e-10;
 Matches 85; Conservative 47; Mismatches 120; Indels 50; Gaps 10;

QY 18 VELPTCMYRLPNV-----HRSYGPAPAGHVOEESNLSIQALESRODDILKRYEL 69
 DB 13 IKLPTCMYRLPNVSLADSLASSGSSSTASSTASSTSCDDTASVAA---RQEKVLRQLEL 69
 QY 70 KAAVGLSKMIQTDPADLDVNTIIQADEPTTLTNALDLSVGLGDKYALKDIYINAPNA 129
 DB 70 KAAVGLSKMIQTDPADLDVNTIIQADEPTTLTNALDLSVGLGDKYALKDIYINAPNA 129
 QY 70 KAAVGLSKMIQTDPADLDVNTIIQADEPTTLTNALDLSVGLGDKYALKDIYINAPNA 129
 DB 70 KAAVGLSKMIQTDPADLDVNTIIQADEPTTLTNALDLSVGLGDKYALKDIYINAPNA 129
 QY 130 SPPLSLIVLRLLCEHFRVLSVTHSSVSV-----PENLLKCFGEONKKOPRODYQ 182
 DB 114 FIYALALAKNMRNLYTIDVKTFTSTWADIGPAAREFEANLAKV---PVNPAIPK---- 167
 QY 183 LGFTLLWKNVPTQMKFSIQTCPIEGGNLAFPLFSFGQKHANVATL---IDSVDI 239
 DB 168 ISVTLLWKNVPTQMKFSIQTCPIEGGNLAFPLFSFGQKHANVATL---IDSVDI 239
 QY 240 AIFQKESGSKKAIVFSSNALSALGSPWLAGNELTVADVLMVLSVLOQIGGCVTPVAV 239
 DB 228 CYQLRKNHTKQVAMVRLDLRLKQKQYFGSGQSVADVGVYSSL-----IRMPAVT 280
 QY 300 QR 301
 DB 281 EK 282

RESULT 13
 ID Q8T060 PRELIMINARY; PRT; 313 AA.
 AC Q8T060;
 DT 01-JUN-2002 (TRENBLREL. 21, Created)
 DT 01-MAR-2004 (TRENBLREL. 26, Last sequence update)
 DE 01-MAR-2004 (TRENBLREL. 26, Last annotation update)
 GN Drosophila melanogaster (fruit fly).
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RA Stapleton M., Brokstein P., Hong L., Agbayan A., Carlson J.,
 RA Change M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liso G.,
 RA Miranda A., Mungall C.J., Munoz V., Pacle J., Paragas V., Park S.,
 RA Patel S., Phoumenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.;
 RT Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY069537; AAL39682.2; -;
 DR FlyBase; FBgn0036515; CG12304.
 DR InterPro; IPR004046; GST_Cterm.
 DR InterPro; IPR010987; GST_C_1like.
 DR Pfam; PF00043; GST_C_1.
 FT NON_TER 1
 SQ SEQUENCE 313 AA; 34522 MW; 5AF9C801877F8531 CRC64;

Query Match 13.4%; Score 224; DB 2; Length 313;
 Best Local Similarity 28.1%; Pred. No. 8e-10;
 Matches 85; Conservative 47; Mismatches 120; Indels 50; Gaps 10;
 QY 18 VELPTCMYRLPNV-----HRSYGPAPAGHVOEESNLSIQALESRODDILKRYEL 69
 DB 13 IKLPTCMYRLPNVSLADSLASSGSSSTASSTASSTSCDDTASVAA---RQEKVLRQLEL 69
 QY 70 KAAVGLSKMIQTDPADLDVNTIIQADEPTTLTNALDLSVGLGDKYALKDIYINAPNA 129
 DB 70 KAAVGLSKMIQTDPADLDVNTIIQADEPTTLTNALDLSVGLGDKYALKDIYINAPNA 129
 QY 70 KAAVGLSKMIQTDPADLDVNTIIQADEPTTLTNALDLSVGLGDKYALKDIYINAPNA 129
 DB 70 KAAVGLSKMIQTDPADLDVNTIIQADEPTTLTNALDLSVGLGDKYALKDIYINAPNA 129
 QY 130 SPPLSLIVLRLLCEHFRVLSVTHSSVSV-----PENLLKCFGEONKKOPRODYQ 182
 DB 126 FIYALALAKNMRNLYTIDVKTFTSTWADIGPAAREFEANLAKV---PVNPAIPK---- 179


```

-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC
DR      EMBL: AB003530; AAF49612.1; -.
DR      InAct; Q9YUR3; -.
DR      Flybase; FBgn0036515; CG12304.
DR      InterPro; IPR010987; GST_C_like.
DR      InterPro; IPR004046; GST_Cterm.
DR      Pfam; PF00043; GST_C; 1.
KW      Protein biosynthesis.
FT      DOMAIN 280
SQ      SEQUENCE 334 AA; 36933 MW; B68FD70AE621990F CRC64;

Query Match      12.9%; Score 214.5; DB 1; Length 334;
Best Local Similarity 25.6%; Pred. No. 5,2e-09;
Matches 85; Conservative 50; Mismatches 120; Indels 77; Gaps 10;

OY      18 VELPTCMTRLPNV---HGRSYGPAPAGHVOESNLSLQA-----54
DB      13 IKLPICMYPLKKNVSLAADSLASGSSSTASSTSSCKLEANRIDRTGHNATCALDLDLS 72
OY      55 -----LESRODDILKRLVELKAAVDELSCMIQTPADLDVTNIQADEPT 99
DB      73 LGRQIQRLIKDDPTASVAPARQEKVKOLELKAQLGQIR-----AGLGVCG-KTFQHT 123
OY      100 TLTVALDLINSVLAGDYGALKDIVINANPASPPLSLVLAHRLCEHFRVLSVHTHSSVK 159
DB      124 TAFQNG-----GLKEVPLQDVINGHNPIFYALALAKNMRNLTYTIDVKTFTHSMA 176
OY      160 SV-----PENLLKCFGEQNKKKQPRQDYQLGFTLIWKNVPKTQMKFSIQMCPLEGEGN 212
DB      177 DIGPARAREPEANLAKV--FVNPALPK---ISVTLWKNCERTHEMISPTMYVPIYGEVN 230
OY      213 IARPLFSLFGQGHANVATL---ISWDVIALFQLEKSSSEKAVERSMNSALGKSPWL 269
DB      231 IIRYIGRVGPAPERYEGSPLCNEIDVLIDICYQLRCNTHKTQVAMVRLDKRLQKQYF 290
OY      270 AGNELTVADVAVLMSVLAQIGGCVTPANVOR 301
DB      291 GGSQMSVADVGVYSSL-----IRMPATVTEK 315

```

Search completed: February 23, 2005, 13:57:11
 Job time : 62.9811 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:25:39 ; Search time 70.371 Seconds
(without alignments)
1758.725 Million cell updates/sec

Title: US-10-622-817-3
Perfect score: 1667
Sequence: 1 MPWGVKRVHSGAPLRYEL.....RWRSCENLAPPTALKLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues
Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneeqp19808:*
2: geneeqp19908:*
3: geneeqp20008:*
4: geneeqp20018:*
5: geneeqp20028:*
6: geneeqp20038:*
7: geneeqp20038:*
8: geneeqp20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 1667 | 100.0 | 320 | 7 | AD58613 Human Pro |
| 2 | 1667 | 100.0 | 320 | 8 | ABM80843 Tumour-as |
| 3 | 1630 | 97.8 | 312 | 2 | AAW25776 UTVI prot |
| 4 | 1630 | 97.8 | 312 | 8 | ADR86551 1-312 am |
| 5 | 1378 | 82.7 | 272 | 7 | ADC10204 Human NOV |
| 6 | 1191 | 71.4 | 229 | 8 | ADR86553 84-312 am |
| 7 | 832 | 49.9 | 161 | 1 | ADR86552 1-161 am |
| 8 | 247 | 14.8 | 51 | 5 | ABP01502 Human ORF |
| 9 | 214.5 | 12.9 | 334 | 4 | ABE62468 Drosophila |
| 10 | 181 | 10.9 | 263 | 4 | ABG23964 Novel hum |
| 11 | 125.5 | 7.5 | 716 | 3 | AAQ1411 Arabidops |
| 12 | 125.5 | 7.5 | 719 | 3 | AAQ1410 Arabidops |
| 13 | 125.5 | 7.5 | 748 | 3 | AAQ1409 Arabidops |
| 14 | 104 | 6.2 | 639 | 4 | ABE62218 Drosophila |
| 15 | 103 | 6.2 | 1512 | 4 | AAU04349 Mammalian |
| 16 | 103 | 6.2 | 1512 | 4 | AAU04349 Mammalian |
| 17 | 103 | 6.2 | 1512 | 4 | AAU04349 Mammalian |
| 18 | 102 | 6.1 | 201 | 8 | AD524684 Bacterial |
| 19 | 101 | 6.1 | 559 | 2 | AAI0682 Polymydr |
| 20 | 101 | 6.1 | 826 | 4 | AAI0682 Polymydr |
| 21 | 101 | 6.1 | 826 | 4 | AAI0682 Polymydr |
| 22 | 101 | 6.1 | 826 | 4 | AAU1908 C. glutam |
| 23 | 101 | 6.1 | 833 | 4 | AAI0682 Polymydr |
| 24 | 101 | 6.1 | 833 | 4 | AAI0682 Polymydr |
| 25 | 101 | 6.1 | 833 | 4 | AAU1907 C. glutam |

| | | | | | |
|----|-------|-----|------|---|--------------------|
| 26 | 101 | 6.1 | 1221 | 4 | AAQ1409 C. glutam |
| 27 | 101 | 6.1 | 1221 | 5 | ABG80321 C. glutam |
| 28 | 101 | 6.1 | 1221 | 5 | AAE18908 Coryneb |
| 29 | 101 | 6.1 | 1221 | 7 | ADD13589 C. glutam |
| 30 | 100.5 | 6.0 | 1200 | 6 | ABU26059 Protein e |
| 31 | 99.5 | 6.0 | 650 | 2 | AAI29614 Mouse PIA |
| 32 | 99.5 | 6.0 | 651 | 2 | AAI29621 Mouse PIA |
| 33 | 97.5 | 5.8 | 217 | 7 | AD563092 Rat Prote |
| 34 | 97.5 | 5.8 | 217 | 7 | AD563092 Rat Prote |
| 35 | 97.5 | 5.8 | 217 | 7 | AD563092 Rat Prote |
| 36 | 97.5 | 5.8 | 217 | 7 | AD563092 Rat Prote |
| 37 | 97.5 | 5.8 | 217 | 7 | AD563092 Rat Prote |
| 38 | 97.5 | 5.8 | 217 | 7 | AD563092 Rat Prote |
| 39 | 97.5 | 5.8 | 218 | 2 | AAI20033 Rat brain |
| 40 | 97.5 | 5.8 | 218 | 2 | AAI20033 Rat brain |
| 41 | 97.5 | 5.8 | 652 | 5 | ABP41869 Human ova |
| 42 | 97 | 5.8 | 559 | 3 | AAI77178 S. venezu |
| 43 | 97 | 5.8 | 2221 | 4 | ABE61339 Streptomy |
| 44 | 97 | 5.8 | 5215 | 8 | ADL91930 Photorhab |
| 45 | 96.5 | 5.8 | 1049 | 6 | ABM67090 Photorhab |

ALIGNMENTS

| | | |
|----------|---|----------------------------|
| RESULT 1 | AD58613 | standard; protein; 320 AA. |
| ID | AD58613 | |
| AC | AD58613 | |
| DT | 29-JAN-2004 | (first entry) |
| DE | Human Protein Q13155, SEQ ID NO 4489. | |
| XX | Human; pain; neuronal tissue; gene therapy; | |
| XX | spinal segmental nerve injury; chronic constriction injury; CCI; | |
| XX | spared nerve injury; SN1; Chung. | |
| OS | Homo sapiens. | |
| XX | WO2003016475-A2. | |
| XX | 27-FEB-2003. | |
| XX | 14-AUG-2002; 2002WO-US025765. | |
| XX | 14-AUG-2001; 2001US-0312147P. | |
| XX | 01-NOV-2001; 2001US-0346382P. | |
| XX | 26-NOV-2001; 2001US-033347P. | |
| XX | (GENO) GEN HOSPITAL CORP. | |
| XX | (FARB) BAYER AG. | |
| XX | Woolf C, D'ursio D, Befort K, Costigan M; | |
| XX | WPI, 2003-268312/26. | |
| XX | GENBANK; Q13155. | |
| XX | New composition comprising two or more isolated polypeptides, useful for | |
| XX | preparing a medicament for treating pain in an animal. | |
| XX | Claim 1, Page; 1017pp; English. | |
| XX | The invention discloses a composition comprising two or more isolated rat | |
| XX | or human polynucleotides or a polynucleotide which represents a fragment, | |
| XX | derivative or allelic variation of the nucleic acid sequence. Also | |
| XX | claimed are a vector comprising the novel polynucleotide, a host cell | |
| XX | comprising the vector, a method for identifying a nucleotide sequence | |
| XX | which is differentially regulated in an animal subjected to pain and a | |
| XX | kit to perform the method, an array, a method for identifying an agent | |
| XX | that increases or decreases the expression of the polynucleotide sequence | |
| XX | that is differentially expressed in neuronal tissue of a first animal | |

CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (SNI) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 320 AA;

Query Match 100.0%; Score 1667; DB 7; Length 320;
 Best Local Similarity 100.0%; Pred. No. 5.4e-167;
 Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPWYQVYPYHGGAPLRLVELPTCMYRLPNVHGRSYGPAAGHVOEESNLSLOLESROD 60
 DB 1 MPWYQVYPYHGGAPLRLVELPTCMYRLPNVHGRSYGPAAGHVOEESNLSLOLESROD 60
 QY 61 DILKRLYEIKAAVVDLSKMTIQPDADLDVTNIIQADEPTTLTTNALDLSVLTGKDYGALK 120
 DB 61 DILKRLYEIKAAVVDLSKMTIQPDADLDVTNIIQADEPTTLTTNALDLSVLTGKDYGALK 120
 QY 121 DIVINANPASPPLSLVLRLLCEHFRVLTSTVHTSSVSKVSPENLLKCFEBOKKKOPROD 180
 DB 121 DIVINANPASPPLSLVLRLLCEHFRVLTSTVHTSSVSKVSPENLLKCFEBOKKKOPROD 180
 QY 181 YOLGFTLLTWKNVPTKQKFSIQTCPIEGEGNIAFLFSLFGQKHNAVNATLIDSWVDIA 240
 DB 181 YOLGFTLLTWKNVPTKQKFSIQTCPIEGEGNIAFLFSLFGQKHNAVNATLIDSWVDIA 240
 QY 241 IFOLKEGSSKEKAIVFRSMNSALGKSPWLAGNELTVADVLMVSVLQOIGGCSVTYPANVO 300
 DB 241 IFOLKEGSSKEKAIVFRSMNSALGKSPWLAGNELTVADVLMVSVLQOIGGCSVTYPANVO 300
 QY 301 RWMRSCEINLAPFNLTALKLK 320
 DB 301 RWMRSCEINLAPFNLTALKLK 320

RESULT 2

ABM80843
 ID ABM80843 standard; protein; 320 AA.

XX
 AC ABM80843;

DT 18-NOV-2004 (first entry)

DE Tumour-associated antigenic target (TAT) polypeptide PRO81501, SEQ:2179.

XX Tumour-associated antigenic target; TAT; human; overexpression; cancer;
 KW tumour; diagnosis; cell proliferative disorder; breast cancer;
 KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
 KW cervical cancer; system cancer; bladder cancer; pancreatic cancer;
 KW chromosomal identification; leukaemia; hybridisation probe;
 KW gene therapy; cytostatic.

XX Homo sapiens.

OS WO2004030615-A2.

XX 15-APR-2004.

XX 29-SEP-2003; 2003WO-US028547.
 PF 02-OCT-2002; 2002US-0414971P.
 XX (GENTECH) GENENTECH INC.
 PA Wu TD, Zhang Z, Zhou Y,
 PI WPI; 2004-347921/32.
 DR N-PSDB; ACN38579.
 XX New tumor-associated antigenic target polypeptides and nucleic acids,
 PT useful in preparing a medicament for treating or detecting a
 PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
 PT prostate cancer or tumor.
 XX Claim 12; SEQ ID NO 2179; 7273bp; English.
 PS The invention relates to human tumour-associated antigenic target (TAT)
 CC polypeptides, and their related nucleic acids. The TAT polypeptides are
 CC overexpressed in cancer tissues compared to normal tissues and may thus
 CC serve as effective targets for the diagnosis and treatment of cancer in
 CC mammals. The invention also relates to nucleic acid and polypeptide
 CC sequences at least 80% identical to the TAT nucleic acid and polypeptide
 CC polypeptides; expression vectors and host cells comprising a TAT nucleic
 CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
 CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
 CC TAT polypeptide; and methods and compositions for the treatment or
 CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
 CC antibodies, antagonists, binding molecules and compositions are useful
 CC for diagnosing or treating a cell proliferative disorder associated with
 CC increased TAT expression, particularly cancers such as breast cancer,
 CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
 CC cancer, pancreatic cancer, cervical cancer, cancers of the central
 CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
 CC used as hybridisation probes, in chromosome and gene mapping, in
 CC chromosome identification and in gene therapy. The present sequence
 CC represents a TAT polypeptide of the invention.

XX Sequence 320 AA;

Query Match 100.0%; Score 1667; DB 8; Length 320;
 Best Local Similarity 100.0%; Pred. No. 5.4e-167;
 Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPWYQVYPYHGGAPLRLVELPTCMYRLPNVHGRSYGPAAGHVOEESNLSLOLESROD 60
 DB 1 MPWYQVYPYHGGAPLRLVELPTCMYRLPNVHGRSYGPAAGHVOEESNLSLOLESROD 60
 QY 61 DILKRLYEIKAAVVDLSKMTIQPDADLDVTNIIQADEPTTLTTNALDLSVLTGKDYGALK 120
 DB 61 DILKRLYEIKAAVVDLSKMTIQPDADLDVTNIIQADEPTTLTTNALDLSVLTGKDYGALK 120
 QY 121 DIVINANPASPPLSLVLRLLCEHFRVLTSTVHTSSVSKVSPENLLKCFEBOKKKOPROD 180
 DB 121 DIVINANPASPPLSLVLRLLCEHFRVLTSTVHTSSVSKVSPENLLKCFEBOKKKOPROD 180
 QY 181 YOLGFTLLTWKNVPTKQKFSIQTCPIEGEGNIAFLFSLFGQKHNAVNATLIDSWVDIA 240
 DB 181 YOLGFTLLTWKNVPTKQKFSIQTCPIEGEGNIAFLFSLFGQKHNAVNATLIDSWVDIA 240
 QY 241 IFOLKEGSSKEKAIVFRSMNSALGKSPWLAGNELTVADVLMVSVLQOIGGCSVTYPANVO 300
 DB 241 IFOLKEGSSKEKAIVFRSMNSALGKSPWLAGNELTVADVLMVSVLQOIGGCSVTYPANVO 300
 QY 301 RWMRSCEINLAPFNLTALKLK 320
 DB 301 RWMRSCEINLAPFNLTALKLK 320

RESULT 3

AAW25776

ID AAW25776 standard; protein; 312 AA.
XX
AC AAW25776;
XX
DT 19-DEC-1997 (first entry)
XX
XX JTV1 protein.
DE JTV1 protein.
XX
XX JTV1; hPMS2; probe; detection; chromosome 7; deletion;
KM mismatch repair gene; hereditary non-polyposis colorectal cancer;
KM homologous recombination.
XX
XX Homo sapiens.
OS
XX MO9708312-A1.
PN 06-MAR-1997.
PD 26-AUG-1996; 96MO-US013598.
XX
XX 24-AUG-1995; 95US-00518862.
PR (UYUO) UNIV JOHNS HOPKINS.
XX
XX Vogelstein B, Kinzler KW, Nicolaides NC;
PI
XX MPI; 1997-179269/16.
DR N-PSDB; AAT86182.
XX
PT Novel chromosome 7 gene, JTV1 - used for detecting chromosome 7
PT deletions, and PMS2 promoter activity.
XX
XX Claim 5; Fig 2; 31pp; English.
PS
XX This sequence is JTV1 protein and is encoded by DNA isolated from human
CC chromosome 7. The JTV1 coding sequence is located upstream from hPMS2.
CC JTV1 cDNA can be used as probes to detect chromosome 7 deletions
CC involving JTV1. Due to the overlapping promoter regions, deletions of
CC JTV1 would also affect PMS2 (a mismatch repair gene) expression, leading
CC to hereditary non-polyposis colorectal cancer. JTV1 can also be used to
CC assay activity or competence of the PMS2 promoter region, the presence of
CC JTV1 suggesting that the PMS2 promoter is intact. JTV1 sequences can also
CC be used to guide homologous recombination at the PMS2 locus
XX
XX Sequence 312 AA;
SQ
Query Match 97.8%; Score 1630; DB 2; Length 312;
Best Local Similarity 100.0%; Pred. No. 4,3e-163;
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 4
ID ADR86551 standard; protein; 312 AA.
XX
AC ADR86551;
XX
DT 18-NOV-2004 (first entry)
XX
XX 1-312 amino acid sequence of p38/JTV-1 protein.
DE
XX p38/JTV-1; Cytostatic; cancer; leukemia; anticancer.
XX
XX Homo sapiens.
OS
XX EPI454628-A2.
PN 08-SEP-2004.
PD 09-SEP-2003; 2003EP-00020344.
XX
XX 03-MAR-2003; 2003KR-00013058.
PR (UYSE-) UNIV SEOUL NAT IND FOUND.
XX
XX Kim S, Park B;
PI
XX MPI; 2004-627822/61.
DR N-PSDB; ADR86548.
XX
PT New isolated p38/JTV-1 protein, useful as medicament for treating cancer
PT e.g., stomach, liver, blood, bone, pancreatic, skin, head or neck cancer
PT and cutaneous or intraocular melanoma, as well as for screening new
PT anticancer agents.
XX
XX Claim 5; SEQ ID NO 4; 47pp; English.
PS
XX The present invention relates to an isolated p38/JTV-1 protein for use as
CC medicament. The p38/JTV-1 protein or the pharmaceutical composition is
CC useful as medicament for treating breast cancer, large intestinal cancer,
CC lung cancer, small cell lung cancer, stomach cancer, liver cancer, blood
CC cancer, bone cancer, pancreatic cancer, skin cancer, head or neck cancer,
CC cutaneous or intraocular melanoma, uterine sarcoma, ovarian cancer,
CC rectal cancer, anal cancer, colon cancer, fallopian tube carcinoma,
CC endometrial carcinoma, cervical cancer, vulval cancer, vaginal carcinoma,
CC Hodgkin's disease, esophageal cancer, small intestine cancer, endocrine
CC cancer, thyroid cancer, parathyroid cancer, adrenal cancer, soft tissue
CC tumour, urethral cancer, penile cancer, prostate cancer, chronic or acute
CC leukemia, lymphocytic lymphoma, bladder cancer, kidney cancer, ureter
CC cancer, renal cell carcinoma, renal pelvic carcinoma, CNS tumour, primary
CC CNS lymphoma, bone marrow tumour, brain stem nerve glioma, pituitary
CC adenoma, or their combination. The protein is useful as a target for
CC screening new anticancer agents. The present sequence represents the 1-
CC 312 amino acid sequence of p38/JTV-1 protein.
XX
XX Sequence 312 AA;
SQ
Query Match 97.8%; Score 1630; DB 8; Length 312;
Best Local Similarity 100.0%; Pred. No. 4,3e-163;
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | |
|----------|-------------|---|
| DE | | 1-161 amino acid sequence of p38/JTV-1 protein. |
| XX | KM | p38/JTV-1; Cytostatic; cancer; leukemia; anticancer. |
| XX | OS | Homo sapiens. |
| XX | PN | EPI454628-A2. |
| XX | PD | 08-SEP-2004. |
| XX | PF | 09-SEP-2003; 2003EP-00020344. |
| XX | PR | 03-MAR-2003; 2003KR-00013058. |
| XX | PA | (UYSE-) UNIV SECUL NAT IND FOUND. |
| XX | PI | Kim S, Park B; |
| XX | DR | WPI; 2004-627822/61. |
| XX | N-PSDB | ADR86549. |
| PT | PT | New isolated p38/JTV-1 protein, useful as medicament for treating cancer e.g., stomach, liver, blood, bone, pancreatic, skin, head or neck cancer |
| PT | PT | and cutaneous or intraocular melanoma, as well as for screening new |
| PT | PT | anticancer agents. |
| PS | PS | Claim 5; SEQ ID NO 5; 47pp; English. |
| XX | XX | The present invention relates to an isolated p38/JTV-1 protein for use as |
| CC | CC | medicament. The p38/JTV-1 protein or the pharmaceutical composition is |
| CC | CC | useful as medicament for treating breast cancer, large intestinal cancer, |
| CC | CC | lung cancer, small cell lung cancer, stomach cancer, liver cancer, blood |
| CC | CC | cancer, bone cancer, pancreatic cancer, skin cancer, head or neck cancer, |
| CC | CC | cutaneous or intraocular melanoma, uterine sarcoma, ovarian cancer, |
| CC | CC | rectal cancer, anal cancer, colon cancer, fallopian tube carcinoma, |
| CC | CC | endometrial carcinoma, cervical cancer, vulval cancer, vaginal carcinoma, |
| CC | CC | Hodgkin's disease, esophageal cancer, small intestine cancer, endocrine |
| CC | CC | cancer, thyroid cancer, parathyroid cancer, adrenal cancer, soft tissue |
| CC | CC | tumour, urethral cancer, penile cancer, prostate cancer, chronic or acute |
| CC | CC | leukemia, lymphocytic lymphoma, bladder cancer, kidney cancer, ureter |
| CC | CC | cancer, renal cell carcinoma, renal pelvic carcinoma, CNS tumour, primary |
| CC | CC | CNS lymphoma, bone marrow tumour, brain stem nerve gliomas, pituitary |
| CC | CC | adenoma, or their combination. The protein is useful as a target for |
| CC | CC | screening new anticancer agents. The present sequence represents the 1- |
| CC | CC | 161 amino acid sequence of p38/JTV-1 protein. |
| XX | XX | |
| SQ | SQ | Sequence 161 AA: |
| | | |
| | | Query Match 49.9%; Score 832; DB 8; Length 161; |
| | | Best Local Similarity 100.0%; Pident. No. 3.9e-79; |
| | | Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |
| QY | | 1 MPWVQVRYHGGAPLRVELPTCMYRLPNVHGSSYGCPAPGAGVQEESNSTLQALESRD 60 |
| DB | | 1 MPWVQVRYHGGAPLRVELPTCMYRLPNVHGSSYGCPAPGAGVQEESNSTLQALESROD 60 |
| QY | | 61 DIKRLTELAAVDGSKMIQTDPADLDVTNIIOADEPTLTITNAIDLNVSGKDYGALK 120 |
| DB | | 61 DIKRLTELAAVDGSKMIQTDPADLDVTNIIOADEPTLTITNAIDLNVSLCKDYALK 120 |
| QY | | 121 DIVINANPASPPLSLVLIRLLCEHFPRVLSVTHTSSVKSV 161 |
| DB | | 121 DIVINANPASPPLSLVLIRLLCEHFPRVLSVTHTSSVKSV 161 |
| | | |
| RESULT 8 | | |
| ID | ABP01502 | standard; protein; 51 AA. |
| XX | ABP01502; | |
| XT | 24-JUN-2002 | (first entry) |
| XX | | |

DE Human ORFX protein sequence SEQ ID NO:2986.

XX Human: open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumor; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypochyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis.

OS Homo sapiens.

PN WO200192523-A2.

PD 06-DEC-2001.

PF 29-MAY-2001; 2001WO-US010836.

PR 30-MAY-2000; 2000US-0206132P.

PR 29-AUG-2000; 2000US-0228716P.

PA (CURA-) CURAGEN CORP.

PI Shinkets RA, Leach MD;

DR WPI, 2002-106308/14.

DR N-PSDB; ABLN17254.

PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
XX hyperproliferative disorders and autoimmune disorders.

PS Disclosure; SEQ ID NO 2986; 1037bp; English.

XX The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-1191 (see Table 1
CC in the specification). ABLN15762 to ABLN27252 encode the human ORFX
CC proteins given in ABLN0010 to ABLN1500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumors, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cyclokin damage. N.B. The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 51 AA;

Query Match 14.8%; Score 247; DB 5; Length 51;
Best Local Similarity 96.1%; Pred. No. 4.5e-16;

Matches 49; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 74 DGLSMIQTTPADLVNTIIQADEPTLTITNALDINSVLGKDYALKDIYI 124
DB 1 DGLSMIQTTPADLVNTIIQADEPTLTITNALDINSVLGKDYALKDIYI 51

RESULT 9
ABR62468

ID ABR62468 standard; protein; 334 AA.

AC ABR62468;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 14196.

KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li FWD, Myers EW,

DR WPI, 2001-656860/75.

DR N-PSDB; ABL06571.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.

PS Disclosure; SEQ ID NO 14196; 21bp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABR57737-
CC ABR72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 334 AA;

Query Match 12.9%; Score 214.5; DB 4; Length 334;
Best Local Similarity 25.6%; Pred. No. 2.8e-13;
Matches 85; Conservative 50; Mismatches 120; Indels 77; Gaps 10;

QY 18 VELPTCMYRLPNV---HRSYGPAPGAGHVOESNLSLQA----- 54
DB 13 IKLPICMPLKNVSLAASLSGSSSTASSTSSCLNRIIDRGKRNATGALDLS 72

QY 55 -----LESRODDILKRLVELKAAVDGSKMIGTPDADLVNTIIQADEPT 99

DB 73 LGRQYRLKDKDTPASVAKQKVEVLKQLELAAQGGQR-----AGLVCQ--KTFOHT 123

QY 100 TTTNALDINSVLGKDYALKDIYINNPASPPLSLVLRLCEHFRVLSVTHHTSSVK 159

DB 124 TAFQNG-----DLKEVPLQDVVINGHPFIPYALLALAKMARNTITIVKTFHSTWA 176

QY 160 SV-----PENLKCGEQNKQKQRPQDYGGLTLKQNVKTKQMFSTQTMCPLEGEGN 212

DB 177 DIGPAAKFEKNAKV--PVNPALPK---ISVTLTKNKEHEMTISSPTMYPIYGEVN 230

QY 213 IARPLFSLFGQKHNAAVNATL---IDSVVDIAIQLKESGKEKAAVFRSMNSALGSPML 269

DB 231 IIRYLGRVGAPEAYRYEGSPCLNEIDLVLDICVLLKCNTHKTVANVRLDLRLQKQYF 290

QY 270 AGNELTVADVVLVMSVLQIGGCSVTVPANVR 301

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: February 23, 2005, 13:36:20 ; Search time 17.81 Seconds
(without alignments)
1341.256 Million cell updates/sec

Title: US-10-622-817-3

Perfect score: 1667

Sequence: 1 MPWGVKPYHGGAPLRYEL.....RWRSCENLAPPTALKLK 320

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfillseq1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--|
| 1 | 1667 | 100.0 | 341 | 4 | US-09-949-016-11312 Sequence 11312, A |
| 2 | 1630 | 97.8 | 312 | 2 | US-08-518-862C-2 Sequence 2, App1 |
| 3 | 103 | 6.2 | 1512 | 3 | US-09-443-184-48 Sequence 48, App1 |
| 4 | 97 | 5.8 | 559 | 2 | US-08-756-317-10 Sequence 4, App1 |
| 5 | 97 | 5.8 | 559 | 4 | US-09-091-609-4 Sequence 2, App1 |
| 6 | 97 | 5.8 | 5215 | 3 | US-09-105-537-2 Sequence 1, App1 |
| 7 | 95 | 5.7 | 559 | 4 | US-09-821-016-1 Sequence 1, App1 |
| 8 | 95 | 5.7 | 559 | 4 | US-10-259-632-1 Sequence 1, App1 |
| 9 | 95 | 5.7 | 559 | 4 | US-10-266-787-1 Sequence 1, App1 |
| 10 | 92 | 5.5 | 2954 | 4 | US-09-150-867-1 Sequence 1, App1 |
| 11 | 91 | 5.5 | 474 | 4 | US-09-248-796A-20321 Sequence 20321, A |
| 12 | 90.5 | 5.4 | 359 | 3 | US-09-540-824-2 Sequence 102, App1 |
| 13 | 90.5 | 5.4 | 597 | 1 | US-08-399-696-102 Sequence 736, App |
| 14 | 88.5 | 5.3 | 1007 | 4 | US-09-538-092-736 Sequence 1178, A |
| 15 | 88 | 5.3 | 427 | 4 | US-09-949-016-11178 Sequence 62, App1 |
| 16 | 88 | 5.3 | 724 | 4 | US-09-900-920-62 Sequence 2, App1 |
| 17 | 88 | 5.3 | 883 | 4 | US-09-976-239-2 Sequence 2, App1 |
| 18 | 88 | 5.3 | 883 | 4 | US-10-289-779B-2 Sequence 4, App1 |
| 19 | 88 | 5.3 | 914 | 4 | US-09-976-239-4 Sequence 4, App1 |
| 20 | 88 | 5.3 | 914 | 4 | US-10-289-779B-4 Sequence 13, App1 |
| 21 | 86.5 | 5.2 | 761 | 3 | US-09-012-710-13 Sequence 13, App1 |
| 22 | 86.5 | 5.2 | 761 | 3 | US-09-556-273-13 Sequence 6834, App |
| 23 | 86 | 5.2 | 211 | 4 | US-09-328-352-6834 Sequence 6, App1 |
| 24 | 86 | 5.2 | 351 | 4 | US-08-178-257-6 Sequence 4904, App |
| 25 | 86 | 5.2 | 471 | 3 | US-09-134-001C-4904 Sequence 18, App1 |
| 26 | 86 | 5.2 | 1365 | 3 | US-09-376-330-18 Patent No. 5194600 |
| 27 | 86 | 5.2 | 1365 | 6 | 5194600-4 |

| | | | | | | |
|----|------|-----|------|---|----------------------|--------------------|
| 28 | 86 | 5.2 | 1365 | 6 | 5194600-4 | Patent No. 5194600 |
| 29 | 85.5 | 5.1 | 1440 | 3 | US-09-357-251-37 | Sequence 37, App1 |
| 30 | 85 | 5.1 | 559 | 4 | US-10-253-509-1 | Sequence 1, App1 |
| 31 | 85 | 5.1 | 1135 | 2 | US-08-469-537A-97 | Sequence 97, App1 |
| 32 | 84.5 | 5.1 | 579 | 4 | US-09-107-433-448 | Sequence 4488, App |
| 33 | 84.5 | 5.1 | 829 | 4 | US-07-670-611-2 | Sequence 2, App1 |
| 34 | 84.5 | 5.1 | 829 | 1 | US-08-220-674-2 | Sequence 2, App1 |
| 35 | 84.5 | 5.1 | 829 | 1 | US-08-445-186-2 | Sequence 2, App1 |
| 36 | 84.5 | 5.1 | 829 | 1 | US-08-446-549-2 | Sequence 2, App1 |
| 37 | 84.5 | 5.1 | 829 | 2 | US-08-446-550-2 | Sequence 2, App1 |
| 38 | 83.5 | 5.0 | 778 | 4 | US-09-583-110-3930 | Sequence 3930, App |
| 39 | 83.5 | 5.0 | 786 | 4 | US-09-107-433-3893 | Sequence 3893, App |
| 40 | 83 | 5.0 | 559 | 2 | US-09-052-339-1 | Sequence 1, App1 |
| 41 | 83 | 5.0 | 559 | 3 | US-09-385-742B-1 | Sequence 1, App1 |
| 42 | 83 | 5.0 | 559 | 4 | US-09-989-786-1 | Sequence 1, App1 |
| 43 | 83 | 5.0 | 767 | 3 | US-08-836-567-8 | Sequence 8, App1 |
| 44 | 83 | 5.0 | 767 | 4 | US-09-606-304-8 | Sequence 8, App1 |
| 45 | 82.5 | 4.9 | 222 | 4 | US-09-248-796A-18146 | Sequence 18146, A |

ALIGNMENTS

```
RESULT 1
US-09-949-016-11312
; Sequence 11312, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11312
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11312

Query Match      100.0%; Score 1667; DB 4; Length 341;
Best Local Similarity 100.0%; Pred. No. 2.4e-181;
Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MPWGVKPYHGGAPLRYELPTCMYRLPVNHGRSYGPAFGAHVOEESNLISQALESSROD 60
        |||||||
DB      22 MPWGVKPYHGGAPLRYELPTCMYRLPVNHGRSYGPAFGAHVOEESNLISQALESSROD 81
        |||||||

QY      61 DILKLYELKAAVNDGSKVIQTPDDADLVNTIIQADEPTTLTNALDLSVGVKQVGA 120
        |||||||
DB      82 DILKLYELKAAVNDGSKVIQTPDDADLVNTIIQADEPTTLTNALDLSVGVKQVGA 141
        |||||||

QY      121 DIVINANPASPLSLVLRLLCEHFRVYSTVHTSSVSVSPENLLKCGEONKQOPROD 180
        |||||||
DB      142 DIVINANPASPLSLVLRLLCEHFRVYSTVHTSSVSVSPENLLKCGEONKQOPROD 201
        |||||||

QY      181 YOLGFTLIWKVNPXTQMKFSIQTMCPIDEGEINIAFLFSLFQKNAVANATLIDSVWDIA 240
        |||||||
DB      202 YOLGFTLIWKVNPXTQMKFSIQTMCPIDEGEINIAFLFSLFQKNAVANATLIDSVWDIA 261
        |||||||

QY      241 IFOLGEGSKKEAAVFRSNNSALGKSPMLAGNELTVADVVLVLSVLOQTGGCVTVPANVO 300
        |||||||
DB      262 IFOLGEGSKKEAAVFRSNNSALGKSPMLAGNELTVADVVLVLSVLOQTGGCVTVPANVO 321
        |||||||

QY      301 RWRSCENLAPPTALKLK 320
```

Db 322 RMWRSCEMLAPFNTALTKLK 341

RESULT 2

US-08-518-862C-2
; Sequence 2, Application US/0851862C
; Patent No. 5843757
; GENERAL INFORMATION:
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth W.
; APPLICANT: Nicolaides, Nicholas C.
; TITLE OF INVENTION: Human JTV1 Gene Overlaps PMS2 Gene
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 1001 G Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/518, 862C
; FILING DATE: 24-AUG-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.49697
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9239
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 312 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-518-862C-2

Query Match 97.8%; Score 1630; DB 2; Length 312;
Best Local Similarity 100.0%; Pred. No. 3.5e-177;
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MPWYQVYPHYGGGAPLRVEIPTCMYRLPVNHGRSYGPAPAGHVOEESLSTQALESRD 60
1 MPWYQVYPHYGGGAPLRVEIPTCMYRLPVNHGRSYGPAPAGHVOEESLSTQALESRD 60
61 DIKRIYELKAAVDGSKMIOPTPDADLDYNTIIQADEPTTLTTNALDINSVLGKDYALK 120
61 DIKRIYELKAAVDGSKMIOPTPDADLDYNTIIQADEPTTLTTNALDINSVLGKDYALK 120
121 DIVYNNPASPPLSLVLRLLCEHFRVLTSTVHTSSVSVSEBENLLKGFGEONKKOPROD 180
121 DIVYNNPASPPLSLVLRLLCEHFRVLTSTVHTSSVSVSEBENLLKGFGEONKKOPROD 180
181 YOLGFTLLIKVNPCKQMKFSIQTMCPISGEGNIAFLFSLFGQKHNANVATLIDSWDIA 240
181 YOLGFTLLIKVNPCKQMKFSIQTMCPISGEGNIAFLFSLFGQKHNANVATLIDSWDIA 240
241 IFOLKEGSSKEKAAVFRSNNLSALGKSPWLAGNELTVADVLMSTLQOIGGCSYTVPANQ 300
241 IFOLKEGSSKEKAAVFRSNNLSALGKSPWLAGNELTVADVLMSTLQOIGGCSYTVPANQ 300
301 RMWRSCEMLAPF 312
301 RMWRSCEMLAPF 312

RESULT 3

US-09-443-184-48
; Sequence 48, Application US/09443184A
; Patent No. 6372431
; GENERAL INFORMATION:
; APPLICANT: Cunningham, Mary Jane
; APPLICANT: Zweigert, Gary
; APPLICANT: Kaseer, Matthew R.
; APPLICANT: Panzer, Scott
; APPLICANT: Selhammer, Jeffrey J.
; APPLICANT: Yue, Henry
; APPLICANT: Baughn, Mariah
; APPLICANT: Azimzai, Yalda
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: MAMMALIAN TOXICOLOGICAL RESPONSE MARKERS
; FILE REFERENCE: PC-0007 US
; CURRENT APPLICATION NUMBER: US/09/443,184A
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 48
; LENGTH: 1512
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6372431 2302721CD1
US-09-443-184-48

Query Match 6.2%; Score 103; DB 3; Length 1512;
Best Local Similarity 22.1%; Pred. No. 0.12;
Matches 45; Conservative 30; Mismatches 69; Indels 60; Gaps 9;

Db 124 INANPASPPLSLVLRLLCEHFRVLTSTVHTSSVSVSEBENLLKGFGEONKKOPROD 183
6 LTVNSGDPLGAL---LAVEHVKDVS---SVEGKENILH----- 41
Db 184 GFTLLIKVNPCKQMKFSIQTMCPISGEGNIAFLFSLFGQKHNANVATLIDSWV 237
42 ---VSENVLFITD-----NSILRYLRVATTAAGLYGS--NLMEHTRIDHWL 82
Qy 238 DIAIFOLKEGSSKEKAAVFRSNNLSALGKSPWLAGNELTVADVLMSTLQOIGGCS 292
Db 83 EFSAATKL--SSCDSPTSTINELNHCLSLRTYLVGNSLSLADLCVWATLK--GNAAWQSO 138
Qy 293 --TVTPANQRMWRSCEMLAPFNT 314
Db 139 KKKKAPVHVHRWFGFLAQOAFOS 162

RESULT 4

US-08-756-317-10
; Sequence 10, Application US/08756317
; Patent No. 5849894
; GENERAL INFORMATION:
; APPLICANT: Clemente, Thomas E.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Minsky, Timothy A.
; APPLICANT: Stark, David M.
; TITLE OF INVENTION: Improved Rhodospirillum Rubrum
; TITLE OF INVENTION: Poly-B-Hydroxalkonocate Synthase
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:57:26 ; Search time 50.2443 Seconds
(without alignments)
2084.158 Million call updates/sec

Title: US-10-622-817-3

Perfect score: 1667

Sequence: 1 MPWYGVKPYHGAGAPLRYVEL.....RWNRSCENLAPNTALXLLK 320

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 1630 | 97.8 | 312 | 16 | US-10-463-676-4 |
| 2 | 1191 | 71.4 | 229 | 16 | US-10-463-676-6 |
| 3 | 832 | 49.9 | 161 | 16 | US-10-463-676-5 |
| 4 | 102 | 6.1 | 201 | 15 | US-10-369-493-13717 |
| 5 | 101 | 6.1 | 826 | 10 | US-09-746-660A-92 |
| 6 | 101 | 6.1 | 833 | 10 | US-09-746-660A-90 |
| 7 | 101 | 6.1 | 1221 | 9 | US-09-919-891-2 |
| 8 | 101 | 6.1 | 1221 | 9 | US-09-738-626-5163 |
| 9 | 101 | 6.1 | 1221 | 15 | US-10-450-055-2 |
| 10 | 101 | 6.1 | 1221 | 17 | US-10-494-675-156 |
| 11 | 100.5 | 6.0 | 687 | 15 | US-10-425-114-58897 |
| 12 | 100.5 | 6.0 | 1200 | 15 | US-10-283-122A-53993 |
| 13 | 97.5 | 5.8 | 652 | 15 | US-10-264-049-3001 |

| | | | | | | |
|----|------|-----|------|----|----------------------|--------------------|
| 14 | 97.5 | 5.8 | 711 | 16 | US-10-437-963-175483 | Sequence 175483, |
| 15 | 97 | 5.8 | 559 | 9 | US-09-364-847-21 | Sequence 21, Appl1 |
| 16 | 97 | 5.8 | 856 | 9 | US-09-364-847-33 | Sequence 33, Appl1 |
| 17 | 97 | 5.8 | 856 | 9 | US-09-364-847-35 | Sequence 35, Appl1 |
| 18 | 97 | 5.8 | 5215 | 9 | US-09-861-289-2 | Sequence 2, Appl1 |
| 19 | 97 | 5.8 | 5215 | 9 | US-09-860-846-2 | Sequence 2, Appl1 |
| 20 | 97 | 5.8 | 5215 | 10 | US-09-988-384B-2 | Sequence 2, Appl1 |
| 21 | 97 | 5.8 | 5215 | 10 | US-09-836-821-2 | Sequence 2, Appl1 |
| 22 | 97 | 5.8 | 5215 | 14 | US-10-271-889-45 | Sequence 45, Appl1 |
| 23 | 96.5 | 5.8 | 914 | 16 | US-10-437-963-153870 | Sequence 153870, |
| 24 | 96.5 | 5.8 | 1148 | 16 | US-10-655-799-29 | Sequence 29, Appl1 |
| 25 | 96 | 5.8 | 606 | 16 | US-10-437-963-115189 | Sequence 115189, |
| 26 | 95.5 | 5.7 | 455 | 15 | US-10-425-114-42633 | Sequence 42633, A |
| 27 | 95 | 5.7 | 559 | 9 | US-09-821-016-1 | Sequence 1, Appl1 |
| 28 | 95 | 5.7 | 559 | 9 | US-09-820-952A-1 | Sequence 1, Appl1 |
| 29 | 95 | 5.7 | 559 | 14 | US-10-218-519-1 | Sequence 1, Appl1 |
| 30 | 95 | 5.7 | 559 | 14 | US-10-259-632-1 | Sequence 1, Appl1 |
| 31 | 95 | 5.7 | 559 | 14 | US-10-266-787-1 | Sequence 1, Appl1 |
| 32 | 95 | 5.7 | 559 | 14 | US-10-252-518-1 | Sequence 1, Appl1 |
| 33 | 95 | 5.7 | 559 | 17 | US-10-914-244-1 | Sequence 1, Appl1 |
| 34 | 95 | 5.7 | 722 | 15 | US-10-369-483-10013 | Sequence 10013, A |
| 35 | 95 | 5.7 | 936 | 14 | US-10-032-585-7442 | Sequence 7442, Ap |
| 36 | 94.5 | 5.7 | 955 | 15 | US-10-282-112A-69912 | Sequence 69912, A |
| 37 | 94 | 5.6 | 203 | 15 | US-10-369-493-8604 | Sequence 8604, Ap |
| 38 | 94 | 5.6 | 468 | 15 | US-10-424-599-260388 | Sequence 260388, |
| 39 | 94 | 5.6 | 1440 | 16 | US-10-437-963-136277 | Sequence 136277, |
| 40 | 93.5 | 5.6 | 569 | 16 | US-10-437-963-123904 | Sequence 123904, |
| 41 | 93.5 | 5.6 | 733 | 15 | US-10-424-599-243766 | Sequence 243766, |
| 42 | 92 | 5.5 | 739 | 15 | US-10-424-599-251522 | Sequence 251522, |
| 43 | 92 | 5.5 | 1206 | 15 | US-10-282-122A-63838 | Sequence 63838, A |
| 44 | 91.5 | 5.5 | 1149 | 16 | US-10-655-799-20 | Sequence 30, Appl1 |
| 45 | 91.5 | 5.5 | 1192 | 15 | US-10-282-122A-64644 | Sequence 64644, A |

ALIGNMENTS

RESULT 1
US-10-463-676-4
; Sequence 4, Application US/10463676
; Publication No. US20040175375A1
; GENERAL INFORMATION:
; APPLICANT: Park, Bum-Joon
; TITLE OF INVENTION: Method for Treating Cancer Using P38/JTV-1 and Method
; TITLE OF INVENTION: for Screening Pharmaceutical Composition for Treating Cancer
; FILE REFERENCE: 012679-091
; CURRENT APPLICATION NUMBER: US/10/463, 676
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: KR 10-2003-13058
; PRIOR FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)...(312)
; OTHER INFORMATION: 1-312 amino acid sequence of p38/JTV-1
US-10-463-676-4
Query Match 97.8%; Score 1630, DB 16; Length 312;
Best Local Similarity 100.0%; Pred. No. 6.4e-153;
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPWYGVKPYHGAGAPLRYVELPTCMRYLRYVHRSYGPAGAGHVOESNLSTQALSSRD 60
DB 1 MPWYGVKPYHGAGAPLRYVELPTCMRYLRYVHRSYGPAGAGHVOESNLSTQALSSRD 60
QY 61 DLKLRYELKAVDGLSKMIGTPDDLDVNTIIOADEPTTLTNALDLNLSVLGKDYALK 120

Db 61 DILKRLYLKAAVDSGSKMIQTPEADLDVTNIIQADEPTTLTNALDLSVGLKDYGALK 120
QY 121 DIVINANPASPPLSLIVLHRLCEHFRVLSVTHSSVSVENLLKCGEONKKQPRD 180
Db 121 DIVINANPASPPLSLIVLHRLCEHFRVLSVTHSSVSVENLLKCGEONKKQPRD 180
QY 181 YOLGFTLWKVNPKTOMKFSIQTMCPFIEGEGNIAFLFSLFGQKNAVANATLIDSWVDIA 240
Db 181 YOLGFTLWKVNPKTOMKFSIQTMCPFIEGEGNIAFLFSLFGQKNAVANATLIDSWVDIA 240
QY 241 IFOLKEGSSKEKAAYFRSMNSALGKSPMLAGNELTVADVILMSVLCQIGCSVTVPANQ 300
Db 241 IFOLKEGSSKEKAAYFRSMNSALGKSPMLAGNELTVADVILMSVLCQIGCSVTVPANQ 300
QY 301 RWRSCENLAPF 312
Db 301 RWRSCENLAPF 312

RESULT 2

US-10-463-676-6
; Sequence 6, Application US/10463676
; Publication No. US20040175375A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Sung-hoon
; APPLICANT: Park, Bum-Joon
; TITLE OF INVENTION: Method for Treating Cancer Using P38/JTV-1 and Method
; TITLE OF INVENTION: for Screening Pharmaceutical Composition for Treating Cancer
; FILE REFERENCE: 012679-091
; CURRENT APPLICATION NUMBER: US/10/463,676
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: KR 10-2003-13058
; PRIOR FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)...(229)
; OTHER INFORMATION: 84-312 amino acid sequence p38/JTV-1
US-10-463-676-6

Query Match 71.4%; Score 1191; DB 16; Length 229;
Best Local Similarity 100.0%; Pred. No. 1.5e-109; Indels 0; Gaps 0;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 DADLDVTNIIQADEPTTLTNALDLSVGLKDYGALKDIVINANPASPPLSLIVLHRLLC 143
Db 1 DADLDVTNIIQADEPTTLTNALDLSVGLKDYGALKDIVINANPASPPLSLIVLHRLLC 60
QY 144 EHFRLVLSVTHSSVSVENLLKCGEONKKQPRDYOGLTLLWKVNPKTOMKFSIQT 203
Db 61 EHFRLVLSVTHSSVSVENLLKCGEONKKQPRDYOGLTLLWKVNPKTOMKFSIQT 120
QY 204 MCPISGEGNIAFLFSLFGQKNAVANATLIDSWVDIAIFOLKEGSSKEKAAYFRSMNSAL 263
Db 121 MCPISGEGNIAFLFSLFGQKNAVANATLIDSWVDIAIFOLKEGSSKEKAAYFRSMNSAL 180
QY 264 GKSFWIAGNELTVADVILMSVLCQIGCSVTVPANVQWRMRSCEMLAPF 312
Db 161 GKSFWIAGNELTVADVILMSVLCQIGCSVTVPANVQWRMRSCEMLAPF 229

RESULT 3

US-10-463-676-5
; Sequence 5, Application US/10463676
; Publication No. US20040175375A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Sung-hoon
; APPLICANT: Park, Bum-Joon

; TITLE OF INVENTION: Method for Treating Cancer Using P38/JTV-1 and Method
; TITLE OF INVENTION: for Screening Pharmaceutical Composition for Treating Cancer
; FILE REFERENCE: 012679-091
; CURRENT APPLICATION NUMBER: US/10/463,676
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: KR 10-2003-13058
; PRIOR FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)...(161)
; OTHER INFORMATION: 1-161 amino acid sequence of p38/JTV-1
US-10-463-676-5

Query Match 49.9%; Score 832; DB 16; Length 161;
Best Local Similarity 100.0%; Pred. No. 3.9e-74; Indels 0; Gaps 0;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPMYQVKPYHGGAPLRYELPTCMYRLPNVHGRSYGAPAGAHVQESNLSIQALESRD 60
Db 1 MPMYQVKPYHGGAPLRYELPTCMYRLPNVHGRSYGAPAGAHVQESNLSIQALESRD 60
QY 61 DILKRLYLKAAVDSGSKMIQTPEADLDVTNIIQADEPTTLTNALDLSVGLKDYGALK 120
Db 61 DILKRLYLKAAVDSGSKMIQTPEADLDVTNIIQADEPTTLTNALDLSVGLKDYGALK 120
QY 121 DIVINANPASPPLSLIVLHRLCEHFRVLSVTHSSVSV 161
Db 121 DIVINANPASPPLSLIVLHRLCEHFRVLSVTHSSVSV 161

RESULT 4

US-10-369-493-13717
; Sequence 13717, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13717
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Pseudomonas fluorescens
US-10-369-493-13717

Query Match 6.1%; Score 102; DB 15; Length 201;
Best Local Similarity 30.5%; Pred. No. 0.17; Indels 22; Gaps 3;
Matches 32; Conservative 15; Mismatches 36; Indels 22; Gaps 3;

QY 211 GNIAFLFSLFGQKNAVNA-TLIDSWVDIAIFOLKEGSSKEKAAYFRSMNSALGKSPWL 269
Db 106 GPAARLITVGAFAFNEEVITRAHNFVKV-----MDLEIGKTPYL 146
QY 270 AGNELTVADVILMSVLCQIGCSVTVPANVQWRMRSCEMLAPF 312
Db 147 AGTEFTIADVSAVYIAHAPBGVSLDDYANVRAWLSTIEALPGF 191

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:34:50 ; Search time 13.3213 Seconds
(without alignments)
2311.294 Million cell updates/sec

Title: US-10-622-817-3

Perfect score: 1667

Sequence: 1 MPMYGVKRYHGGAPLRLVEL.....RWRSCENLAPFTALKLK 320

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : PIR_79:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 125.5 | 7.5 | 719 | 2 T52043 | probable glutamate |
| 2 | 111 | 6.7 | 728 | 2 T01200 | probable glutamate |
| 3 | 105.5 | 6.3 | 222 | 2 G82441 | probable glutathio |
| 4 | 97.5 | 5.8 | 218 | 2 A29036 | glutathione transf |
| 5 | 97.5 | 5.8 | 645 | 2 JCS517 | Gu/RNA helicase II |
| 6 | 97 | 5.8 | 559 | 2 A3604 | poly(3-hydroxyalka |
| 7 | 96.5 | 5.8 | 174 | 2 S74241 | limonacyl-cRNA syn |
| 8 | 96 | 5.8 | 1675 | 2 S70770 | hypothetical prote |
| 9 | 94 | 5.6 | 218 | 2 B29231 | glutathione transf |
| 10 | 93 | 5.6 | 393 | 2 D64391 | flavoprotein - Met |
| 11 | 92 | 5.5 | 210 | 2 B85594 | probable transfera |
| 12 | 92 | 5.5 | 210 | 2 P64821 | probable glutathio |
| 13 | 92 | 5.5 | 210 | 2 P90743 | probable transfera |
| 14 | 92 | 5.5 | 744 | 2 AE2703 | hypothetical prote |
| 15 | 92 | 5.5 | 764 | 2 E97485 | (P)PPGP synthetas |
| 16 | 92 | 5.5 | 867 | 2 S72842 | methionine synthas |
| 17 | 92 | 5.5 | 1036 | 2 AF2338 | two-component sens |
| 18 | 92 | 5.5 | 1306 | 2 E87072 | hypothetical prote |
| 19 | 92 | 5.5 | 2954 | 2 T1456 | kinasin-related pr |
| 20 | 91.5 | 5.5 | 218 | 1 XURIG4 | glutathione transf |
| 21 | 91.5 | 5.5 | 1192 | 2 G70513 | 5-methyltetrahydro |
| 22 | 91 | 5.5 | 2279 | 2 T42531 | acetyl-CoA carboxy |
| 23 | 91 | 5.5 | 2280 | 2 T38906 | ras-associated pro |
| 24 | 90.5 | 5.4 | 359 | 2 T37921 | probable membrane |
| 25 | 90 | 5.4 | 710 | 2 S67098 | glutathione transf |
| 26 | 89.5 | 5.4 | 218 | 2 B34159 | killer toxin resis |
| 27 | 89.5 | 5.4 | 1365 | 1 BVBK5 | hypothetical prote |
| 28 | 89 | 5.3 | 2480 | 2 D84904 | glutathione transf |
| 29 | 88.5 | 5.3 | 209 | 1 XUFPI1 | glutathione transf |

| | | | | | |
|----|------|-----|------|----------|--------------------|
| 30 | 88.5 | 5.3 | 1007 | 2 S48535 | rho-type GTPase-ac |
| 31 | 88 | 5.3 | 207 | 2 D72272 | conserved hypothe |
| 32 | 88 | 5.3 | 218 | 2 D26946 | glutathione transf |
| 33 | 88 | 5.3 | 437 | 2 I51238 | translation elonga |
| 34 | 87.5 | 5.2 | 214 | 2 B4681 | glutathione transf |
| 35 | 87.5 | 5.2 | 1065 | 2 A43421 | SEC8 protein - yea |
| 36 | 87.5 | 5.2 | 1292 | 2 T31462 | probable magnesium |
| 37 | 87.5 | 5.2 | 1714 | 1 S18644 | multifunctional am |
| 38 | 87.5 | 5.2 | 2586 | 2 T21676 | hypothetical prote |
| 39 | 87 | 5.2 | 586 | 2 C71695 | ctp synthase (pyr |
| 40 | 86.5 | 5.2 | 208 | 1 S43851 | glutathione transf |
| 41 | 86.5 | 5.2 | 379 | 2 D97198 | methy1-accepting c |
| 42 | 86.5 | 5.2 | 436 | 2 I51237 | translation elonga |
| 43 | 86.5 | 5.2 | 439 | 2 D86440 | unknown protein [1 |
| 44 | 86.5 | 5.2 | 749 | 2 A57276 | MAO1 protein - yea |
| 45 | 86 | 5.2 | 377 | 2 P69172 | hypothetical prote |

ALIGNMENTS

RESULT 1

T52043 Probable glutamate-cRNA ligase (EC 6.1.1.17) [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #ext_change 09-Jul-2004

C/Accession: T52043

R/Day, I.S.; Golovkin, M.; Reddy, A.S.

Biochim. Biophys. Acta 1399, 219-24, 1998

A/Title: Cloning of the cDNA for glutamyl-cRNA synthetase from Arabidopsis thaliana.

A/Reference number: Z24836; MUID:9765600; PMID:9765600

A/Accession: T52043

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-719 <DAY>

A/Cross-references: UNIPROT:O82462; EMBL:AF067773; PIDD:AA036469.1

A/Suprafamily: Yeast glutamate-cRNA ligase; glutamine-cRNA ligase homology

C/Keywords: ligase

Query Match 7.5%; Score 125.5; DB 2; Length 719;

Best Local Similarity 27.3%; Pred. No. 0.017;

Matches 48; Conservative 26; Mismatches 53; Indels 49; Gaps 9;

QY 128 PASPPLSLVLRLLCEHFRVSLVTHSSVSPENLLKCFGEQKQKQPRODYGFTL 187

DB 10 PESPPLSVVALSLASPV---TIDSSMAATTVSFFV---PSDGRKLN-----GATV 55

QY 188 IKKNVPTCKMPSIQTMCPIDEGNIAFLFSLFGQKHNAVNATLIDSVVDIA-IFOLKE 246

DB 56 LRLRY-----GRSAKLLPFFYG--NNAPSSQIDENVVYASVF--SS 93

QY 247 GSSKEKAIVFRSMNSALGKSPWLAGNELTVADVVLMSVLOQIGGCVTPVPAVQRW 302

DB 94 GSEFENAC--GRVDYLESSTFLVGHSHIADVAIWSALAGTG-----QRW 137

RESULT 2

T01200 Probable glutamate-cRNA ligase (EC 6.1.1.17) F21E10.12 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #ext_change 09-Jul-2004

R/Davieson, S.; Rohlfing, T.; David, M.; O'Brian, D.

submitted to the EMBL Data Library, April 1998

A/Description: The sequence of A. thaliana F21E10.

A/Reference number: Z14258

A/Accession: T01200

A/Status: translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-728 <DAY>

A/Cross-references: UNIPROT:O65253; EMBL:AF058914; NID:G3047074; PID:G3047084; GSPDB:GNO

A/Experimental source: cultivar Columbia

C/Genetics:

A:Gene: ATSP:F21E10.12
A:Map position: 5
A:Introns: 47/2: 89/3; 141/1; 503/3; 659/3
C:Superfamily: yeast glutamate-tRNA ligase; glutamine-tRNA ligase homology
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis
F:223-499/Domain: glutamine-tRNA ligase homology <EGL>

Query Match 6.7%; Score 111; DB 2; Length 728;
Best Local Similarity 25.9%; Pred. No. 0.3;
Matches 48; Conservative 26; Mismatches 53; Indels 58; Gaps 10;

QY 128 PASPPISLVLRLLCEHFRVLSTVHTSHSVKSPENILKCPGEONKQRPQDQLGFTL 187
DB 10 PESPPLSVIALSLASPV---TIDSSAAATTVPSPV---FSDGRKAN-----GATV 55
QY 188 IKWNPVKQMKFSIQTMCPIEBEGNIARFLPSFGQKNANVANAT-----LIDSWVD 238
DB 56 LRRVY-----GRSACKLPDPYV--NNAPDSQVSTILCINMKIDEMVD 95
QY 239 IAFOLKESKSKRAVFRSNMSALGKSPWLAGNELTVADVLMVSLQOIGGGSVTVPA 297
DB 96 YASVF--SSGSEFENAC--GRVDKLTBSSTFLVGHSLIADVAIWSALAGTG----- 143
QY 298 NVQRP 302
DB 144 --QRP 146

RESULT 3
G82441
probable glutathione S-transferase VCA0584 [imported] - vibrio cholerae (strain N16961)

C:Species: vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: G82441
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gilm, M.L.; Dodson, R.J.;
Haridison, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Drygiol, I.; Sellers, F.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406: 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:2040683; PMID:10952301
A:Accession: G82441
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-222 <HEI>
A:Cross-references: UNIPROT:Q9RM05; GB:AE004389; GB:AE003853; NID:G9657989; PIDN:AAF9648
A:Experimental source: serogroup O1, strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0584
A:Map position: 2
C:Superfamily: hypothetical protein b2302

Query Match 6.3%; Score 105.5; DB 2; Length 222;
Best Local Similarity 25.7%; Pred. No. 0.17;
Matches 28; Conservative 22; Mismatches 38; Indels 21; Gaps 4;

QY 197 MKFSIQTMCPIEBEGNIARFLPSFGQKNANVANATLIDSWVDIAFOLKESKSKRAV 256
DB 97 LMFQMSGVPRMGOANV---FYRYPEKIQP-----IDRY-----QKEGRRLF 137
QY 257 RSNMSALGKSPWLAGNELTVADVLMVSLQOIGGGSVTVPAVQRM 303
DB 138 EVMDDQLAQNPLYAGDEYTIADTFPWRIRHMGISIDGLTHQRNM 186

RESULT 4
A29036
glutathione transferase (EC 2.5.1.18) Yb3 - rat

C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C:Accession: A29036
R:Abrahamovitz, M.; Lisowsky, I.
J. Biol. Chem. 262: 7770-7773, 1987
A:Title: Selective expression of a unique glutathione S-transferase Yb3 gene in rat brain

A:Reference number: A29036; MUID:87222405; PMID:3584141
A:Accession: A29036
A:Molecule type: mRNA
A:Residues: 1-218 <ABR>
A:Cross-references: UNIPROT:P08009; GB:J02744; NID:G204512; PIDN:AAA1292.1; PID:G20451
C:Superfamily: glutathione transferase
C:Keywords: transferase

Query Match 5.8%; Score 97.5; DB 2; Length 218;
Best Local Similarity 25.6%; Pred. No. 0.78;
Matches 33; Conservative 24; Mismatches 41; Indels 31; Gaps 5;

QY 179 QDYOLGFTLIWKNVPTQMKFSIQTMCPIEBEGNIARFLPSFGQKNANVANAT----- 231
DB 49 EKFKLG--LDFPNLP-----YLDGSHKITQSMILKRL-----GRKNLCETEEERLRV 97
QY 232 -----LIDSWVDIA-----IFOLKESKSKRAVFRSNMSALGKSPWLAGNELTVAD 278
DB 98 DILENOLDNRNVLARLCYNDPFEKLRGYLEOLPGWRLYSEFLGKRPWPAQDKITFVD 157
QY 279 VVLMGVLQO 287
DB 158 FIAYDVLR 166

RESULT 5
JC5517
Gu/RNA helicase II binding protein - human

C:Species: Homo sapiens (man)
C:Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 05-Nov-1999
C:Accession: JC5517
R:Valdez, B.C.; Henning, D.; Perlaaky, L.; Busch, R.K.; Busch, H.
Biochem. Biophys. Res. Commun. 234, 335-340, 1997
A:Title: Cloning and characterization of Gu/RN-II binding protein.
A:Reference number: JC5517; MUID:97320420; PMID:9177271
A:Accession: JC5517
A:Molecule type: mRNA
A:Residues: 1-645 <VAL>
A:Cross-references: GB:U78524; NID:g1696006; PIDN:AA858488.1; PID:g1696007
A>Note: It is uncertain whether Met-4 or Met-6 is the initiator
C:Comment: This protein is localized to the nucleus and interacts with Gu/RNA helicase I
C:Keywords: phosphoprotein
F:50-58/Region: nuclear location signal
F:362-374/Region: nuclear location signal
F:514-517, 551-554, 592-595, 606-609/Region: 4-residue repeats (N-T-S-L)
F:717, 732, 412, 432, 460, 461, 462, 470, 614, 627/Binding site: phosphate (Ser) (covalent) #st
F:65, 394, 624/Binding site: phosphate (Thr) (covalent) #stratus predicted

Query Match 5.8%; Score 97.5; DB 2; Length 645;
Best Local Similarity 20.5%; Pred. No. 3.7;
Matches 79; Conservative 56; Mismatches 129; Indels 121; Gaps 20;

QY 16 LRV-ELPTCK-YRLPVNHRSGYPAGAGHVOESNTSLQLESRODDILKELVELKAAV 73
DB 8 LRVSEIQLVLYGVRNKRHGRKHELLTKALHLKAG--CSFAVQMKIKELYRRRFPQK--- 62
QY 74 DGLSKMIOTPPDADIVNTIIQADEPTTLTNMLDINSVIGDYGLKQIVINANPASPPL 133
DB 63 -----IMRP-ADLSIPNHSSPMATLSPS-----TIPQLTYDGHFASPL 102
QY 134 ---SLI-VLRLLCEHF-RVLSTVHTSHSVKSP-----ENILK--CRGEONKQ----- 176
DB 103 LRVSLIGPGRHELEPHLTGALHPVPDILKQLRPLYDLDLILKRTSLASNSQRFRET 162
QY 177 -----PRDYOLGFTLIWKNVPTQMKFSIQ-----TMCPIEGEG----- 211
DB 163 FAPALTPQGVQVQISSM--DISGKCDPTVOVLRFCLSETSCPOEDHFPFPLCVKVT 219
QY 212 -----NIARFLPSFGQKNANVANATLIDSWV-----DI 239
DB 220 KPQSLPGVILPTKNGVEKRRSRPNTISLV-----RISTTVPTIVTSWTAETGRNYSW 274
QY 240 ALFOLKESG-----KEKAVFRSNMSALGKSPWLAGNELTVADVLMVSLQOIGCG 291

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:33:04 ; Search time 60.9145 Seconds
(without alignments)
2694.512 Million cell updates/sec

Title: US-10-622-817-3

Perfect score: 1667

Sequence: 1 MPWGVKPYHGGAPLRLV.....RMWRSCENLAPNTALKLK 320

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 03:*
1: uniprot_gprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1 | 1667 | 100.0 | 320 | 1 | MC62 HUMAN |
| 2 | 1663 | 99.8 | 320 | 2 | Q96CZ5 |
| 3 | 1464 | 87.8 | 320 | 2 | Q8R010 |
| 4 | 1464 | 87.8 | 320 | 2 | Q8R2Y6 |
| 5 | 1444 | 86.6 | 320 | 1 | MC62 CRIGR |
| 6 | 1260 | 75.6 | 280 | 2 | Q8R3V2 |
| 7 | 1144.5 | 68.7 | 311 | 2 | Q6DK86 |
| 8 | 1139.5 | 68.4 | 311 | 2 | Q6IN04 |
| 9 | 1132.5 | 67.9 | 311 | 2 | Q7Z1D7 |
| 10 | 887.5 | 53.2 | 321 | 2 | Q7T3C0 |
| 11 | 267.5 | 16.0 | 340 | 2 | Q7Q7A3 |
| 12 | 224 | 13.4 | 301 | 2 | Q7KUM5 |
| 13 | 224 | 13.4 | 313 | 2 | Q8T060 |
| 14 | 224 | 13.4 | 322 | 2 | Q6NKM4 |
| 15 | 214.5 | 12.9 | 334 | 1 | MC62 DROME |
| 16 | 125.5 | 7.5 | 719 | 2 | Q6S253 |
| 17 | 111 | 6.7 | 728 | 2 | Q6S253 |
| 18 | 106.5 | 6.4 | 880 | 2 | Q8ERT1 |
| 19 | 106.5 | 6.4 | 913 | 2 | Q6D8B3 |
| 20 | 106.5 | 6.4 | 913 | 2 | Q6PRQ4 |
| 21 | 105.5 | 6.3 | 222 | 2 | Q9KM05 |
| 22 | 105 | 6.3 | 240 | 2 | Q8BTR1 |
| 23 | 105 | 6.3 | 1512 | 1 | SYEP MOUSE |
| 24 | 105 | 6.3 | 1800 | 2 | Q7PRA2 |
| 25 | 104.5 | 6.3 | 922 | 2 | Q6T8M5 |
| 26 | 104 | 6.2 | 961 | 2 | Q9VB85 |
| 27 | 104 | 6.2 | 702 | 2 | Q8IHC6 |
| 28 | 104 | 6.2 | 841 | 2 | Q8IMQ2 |
| 29 | 103 | 6.2 | 328 | 2 | Q86X73 |
| 30 | 103 | 6.2 | 869 | 2 | Q6PD57 |
| 31 | 103 | 6.2 | 1213 | 2 | Q8FRD2 |

| | | | | | | |
|----|-------|-----|------|---|------------|--------------------|
| 32 | 102.5 | 6.1 | 372 | 2 | Q87XN7 | Q87XN7 pseudomonas |
| 33 | 102.5 | 6.1 | 715 | 2 | Q9LIZ8 | Q9LIZ8 oryza sativ |
| 34 | 102 | 6.1 | 559 | 2 | Q8K023 | Q8K023 pseudomonas |
| 35 | 102 | 6.1 | 559 | 2 | Q88D25 | Q88D25 pseudomonas |
| 36 | 102 | 6.1 | 1486 | 2 | Q6TXE9 | Q6TXE9 rattus norv |
| 37 | 101 | 6.1 | 1196 | 2 | Q6W580 | Q6W580 corynebacte |
| 38 | 101 | 6.1 | 1221 | 2 | Q8NQD1 | Q8NQD1 corynebacte |
| 39 | 100.5 | 6.0 | 454 | 2 | Q6UDR4 | Q6UDR4 mus musculu |
| 40 | 100.5 | 6.0 | 1201 | 2 | Q6NH83 | Q6NH83 corynebacte |
| 41 | 99.5 | 6.0 | 651 | 1 | PIA1 MOUSE | Q88907 mus musculu |
| 42 | 99 | 5.9 | 203 | 2 | Q6W112 | Q6W112 thizobium s |
| 43 | 98.5 | 5.9 | 222 | 2 | Q8EGM4 | Q8EGM4 shewanella |
| 44 | 98.5 | 5.9 | 582 | 2 | Q6ECB1 | Q6ECB1 yersinia pe |
| 45 | 97.5 | 5.8 | 217 | 1 | GTW3_RAT | P08009 rattus norv |

ALIGNMENTS

RESULT 1
MC62_HUMAN
ID MC62_HUMAN STANDARD; PRT; 320 AA.
AC Q13155; Q9P1L2;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Multisynthetase complex auxiliary component p38 (JTV-1 protein)
DE (PRO0992).
GN Name=JTV1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96115582; PubMed=8666379;
RA Nicolaides N.C., Kinzler K.W., Vogelstein B.;
RT "Analysis of the 5' region of PMS2 reveals heterogeneous transcripts
and a novel overlapping gene";
RL Genomics 29:329-334(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshuler S.F., Zeeberg B., Buellow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.U., Ueda T.B., Toshlyuk S., Carninci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE OF 197-320 FROM N.A.
RC TISSUE=Fetal liver;
RX Zhang C., Yu Y., Zhang S., Wei H., Zhou G., Ouyang S., Luo L., Bi J.,
RA Liu M., He F.;
RT "Functional prediction of the coding sequences of 121 new genes
deduced by analysis of cDNA clones from human fetal liver";
RL submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [4]

RP INTERACTION WITH FUBP1.
 RX MEDLINE=22716800; PubMed=12819782; DOI=10.1038/ng1182;
 RA Kim M.J., Park B.-J., Kang Y.-S., Kim H.J., Park J.-H., Kang J.W.,
 RA Lee S.W., Han J.M., Lee H.-W., Kim S.,
 RT "Downregulation of FUSE-binding protein and c-myc by tRNA synthetase
 RT cofactor p38 is required for lung cell differentiation.";
 RL Nat. Genet. 34:330-336(2003).
 CC -1- FUNCTION: Probable core protein of the multisynthetase complex
 CC that serves as a template for the assembly of the supramolecular
 CC structure. Mediates ubiquitination of FUBP1 and its degradation by
 CC the proteasome.
 CC -1- SUBUNIT: Component of the multisynthetase complex which is
 CC comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the
 CC monospecific isoleucyl, leucyl, glutamyl, methionyl, lysyl,
 CC arginyl, and aspartyl-tRNA synthetases as well as three auxiliary
 CC proteins, p18, p48 and p43. Binds FUBP1.
 CC -1- SIMILARITY: Contains 1 GST-like domain.
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
 CC frameshift in position 312.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, U24169; AAC50391.1; ALT FRAME.
 DR EMBL, BC028853; AAH02853.1; -
 DR EMBL, BC010156; AAH010156.1; -
 DR EMBL, AF116615; AAF71039.1; -
 DR H-InvDB, HIX0006460; -
 DR MIM, 600859; -
 DR InterPro: IPR010987; GST_C-like.
 DR InterPro: IPR004046; GST_Cterm.
 DR Pfam, PF00043; GST_C; 1.
 KW Protein biosynthesis.
 SQ SEQUENCE 320 AA; 35349 MW; F253726B63C12B8B CRC64;
 Query Match 100.0%; Score 1667; DB 1; Length 320;
 Best Local Similarity 100.0%; Pred. No. 66-127;
 Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MPYQVPRYHGGAGPLRVELPTCMYRLPNVHGRSGYPAGAGHYOESNLSIQALESROD 60
 DB 1 MPYQVPRYHGGAGPLRVELPTCMYRLPNVHGRSGYPAGAGHYOESNLSIQALESROD 60
 QY 61 DILKRLYEKAAVDGSKM1QTPDADLDVTNIIQADEPTTLTTNALDINSVLGKDYGALK 120
 DB 61 DILKRLYEKAAVDGSKM1QTPDADLDVTNIIQADEPTTLTTNALDINSVLGKDYGALK 120
 QY 121 DIVINANPASPLSLVLRHLCEHFRVLSVHTHSSVSKVPENLKCFEBOKKKOPROD 180
 DB 121 DIVINANPASPLSLVLRHLCEHFRVLSVHTHSSVSKVPENLKCFEBOKKKOPROD 180
 QY 181 YQUGFTLLIKWNPXTOMKFSIQTCPIEGEGNIARFLFSLFQGNANVATLIDSWDIA 240
 DB 181 YQUGFTLLIKWNPXTOMKFSIQTCPIEGEGNIARFLFSLFQGNANVATLIDSWDIA 240
 QY 241 IFOLKEGSSKEKAIVFRSMNSALGKSPMLAGNELTVADVLMVSVLQOIGGCSVTVPANQ 300
 DB 241 IFOLKEGSSKEKAIVFRSMNSALGKSPMLAGNELTVADVLMVSVLQOIGGCSVTVPANQ 300
 QY 301 RWRSGCENLAPFNLTALKLK 320
 DB 301 RWRSGCENLAPFNLTALKLK 320
 RESULT 2
 ID 096C25 PRELIMINARY; PRT; 320 AA.
 AC 096C25; 096C25;

DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE UT11.
 GN Name=UTV1;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Pelngold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Heife F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshlyuk S., Carninci P., Prange C.,
 RA Bosak S.A., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Richards S., Morley K.C., Hais S., Garcia A.M., Gay L.V., Hulyk S.W.,
 RA Villalón D.K., Muzny K.C., Hais S., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A.C., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalins D.E., Schmech A., Schein J.E.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA Strausberg R.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL, BC013630; AAH13630.1; -
 DR InterPro: IPR004046; GST_Cterm.
 DR InterPro: IPR010987; GST_C-like.
 DR Pfam, PF00043; GST_C; 1.
 SQ SEQUENCE 320 AA; 35335 MW; 19F14BF58612E08 CRC64;
 Query Match 99.8%; Score 1663; DB 2; Length 320;
 Best Local Similarity 99.7%; Pred. No. 1,36-126;
 Matches 319; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MPYQVPRYHGGAGPLRVELPTCMYRLPNVHGRSGYPAGAGHYOESNLSIQALESROD 60
 DB 1 MPYQVPRYHGGAGPLRVELPTCMYRLPNVHGRSGYPAGAGHYOESNLSIQALESROD 60
 QY 61 DILKRLYEKAAVDGSKM1QTPDADLDVTNIIQADEPTTLTTNALDINSVLGKDYGALK 120
 DB 61 DILKRLYEKAAVDGSKM1QTPDADLDVTNIIQADEPTTLTTNALDINSVLGKDYGALK 120
 QY 121 DIVINANPASPLSLVLRHLCEHFRVLSVHTHSSVSKVPENLKCFEBOKKKOPROD 180
 DB 121 DIVINANPASPLSLVLRHLCEHFRVLSVHTHSSVSKVPENLKCFEBOKKKOPROD 180
 QY 181 YQUGFTLLIKWNPXTOMKFSIQTCPIEGEGNIARFLFSLFQGNANVATLIDSWDIA 240
 DB 181 YQUGFTLLIKWNPXTOMKFSIQTCPIEGEGNIARFLFSLFQGNANVATLIDSWDIA 240
 QY 241 IFOLKEGSSKEKAIVFRSMNSALGKSPMLAGNELTVADVLMVSVLQOIGGCSVTVPANQ 300
 DB 241 IFOLKEGSSKEKAIVFRSMNSALGKSPMLAGNELTVADVLMVSVLQOIGGCSVTVPANQ 300
 QY 301 RWRSGCENLAPFNLTALKLK 320
 DB 301 RWRSGCENLAPFNLTALKLK 320

```
RESULT 3
OBR010 PRELIMINARY; PRT; 320 AA.
ID OBR010
AC OBR010;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Jtvl-pending protein.
GN Name=Jtvl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; and FVB/N-3; TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnae.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshlyuk S., Carninci P., Prange C.,
RA Raha S.S., Loguelli N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor; C3;
RA Strausberg R.;
RN Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N-3; TISSUE=Mammary tumor;
RA Strausberg R.;
RN Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024410; AAH24410.1; -.
DR EMBL; BC026972; AAH26972.1; -.
DR MGD; MGI:2385337; Jtvl.
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR010987; GST_C_1like.
DR Pfam; PF00043; GST_C_1.
SQ SEQUENCE 320 AA; 35396 MW; 1745D7EB4BC367D0 CRC64;

Query Match 87.8%; Score 1464; DB 2; Length 320;
Best Local Similarity 87.2%; Pred. No. 1.8e-110;
Matches 279; Conservative 16; Mismatches 25; Indels 0; Gaps 0;
```

```
DB 181 YOLGFTLLWKVNPKTQMKFSPVOTMCPICEGNIARFLFSLFGQKHNAVTLTLLDSVWDIA 240
OY 241 IOLKEGSEKKAAPFRSNNSALGKSPWLAGNELTADVVLMSVLCQIGGCVTPVANYO 300
DB 241 MFOLEBSSKKAAPFRSNNSALGKSPWLVGNELTADVVLMSVLCQIGGSGAFTNVO 300
OY 301 RWRSCENLAPPTALKLK 320
DB 301 RWLKCENLAPPTALQLK 320

RESULT 4
OBR2Y6 PRELIMINARY; PRT; 320 AA.
ID OBR2Y6
AC OBR2Y6;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Jtvl-pending protein.
GN Name=Jtvl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II; TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnae.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshlyuk S., Carninci P., Prange C.,
RA Raha S.S., Loguelli N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II; TISSUE=Mammary tumor;
RA Strausberg R.;
RN Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC026958; AAH26958.1; -.
DR MGD; MGI:2385237; Jtvl.
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR010987; GST_C_1like.
DR Pfam; PF00043; GST_C_1.
SQ SEQUENCE 320 AA; 35423 MW; 1C21F1A74C9882B4 CRC64;

Query Match 87.8%; Score 1464; DB 2; Length 320;
Best Local Similarity 87.2%; Pred. No. 1.8e-110;
Matches 279; Conservative 16; Mismatches 25; Indels 0; Gaps 0;
```

```

Db      121 DIVINANPASPPLSLVLRLLCERFVSLVTHSSVKNVPEMLKCGEGQARKQSRHE 180
Qy      181 YOLGFTLLIMKNVPTQMKFSIQTMCPIDEGNIARFLFSLFGQKNAVANATLIDSWDIA 240
Db      181 YOLGFTLLIMKNVPTQMKFSVQTMCPIDEGNIARFLFSLFGQKNAVANATLIDSWDIA 240
Qy      241 IFOLKEGSSKERAAPVRSNMSALGKSPMLAGNELTVADVLMVSVLQQTGGSSGVAAPVAVQ 300
Db      241 MFOLKEGSSKERAAPVRSNMSALGKSPMLAGNELTVADVLMVSVLQQTGGSSGVAAPVAVQ 300
Qy      301 RMWRSCENLAPPTALTKLK 320
Db      301 RMWRSCENLAPPTALTKLK 320

```

```

RESULT 5
MCA2_CRIGR STANDARD; PRT; 320 AA.
AC 09WVF;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 25-OCT-2004 (Rel. 45, Last annotation update)
OS Multisynthetase complex auxiliary component p38.
OC Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OK NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A..
RC TISSUE=ovary;
RX MEDLINE=99096915; PubMed=9878398; DOI=10.1006/jmbi.1998.2316;
RA Quevillon S., Robinson J.-C., Berthomieu E., Siatecka M., Mirande M.;
RT "Macromolecular assemblage of aminocyl-tRNA synthetases:
RT identification of protein-protein interactions and characterization of
RT a core protein."
RL J. Mol. Biol. 285:183-195(1999).

```

```

CC -1- FUNCTION: Probable core protein of the multisynthetase complex
CC that serves as a template for the assembly of the supramolecular
CC structure.
CC -1- SUBUNIT: Component of the multisynthetase complex which is
CC comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the
CC monospecific isoleucyl, leucyl, glutamyl, methionyl, lysyl,
CC asynyl, and aspartyl-tRNA synthetases as well as three auxiliary
CC proteins, p38, p48 and p43.
CC -1- SIMILARITY: Contains 1 GST-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: A0702727; AAD38422.1; -.
DR InterPro: IPR010987; GST_C_1like.
DR InterPro: IPR004046; GST_C_1like.
DR Pfam: PF00043; GST_C_1.
KW Protein biosynthesis.
SQ SEQUENCE 320 AA; 35433 MW; 6D24ED033ABEC810A CRC64;

```

```

Query Match      86.6%; Score 1444; DB 1; Length 320;
Best Local Similarity 86.6%; Pred. No. 7,6e-109;
Matches 277; Conservative 16; Mismatches 27; Indels 0; Gaps 0;

```

```

Qy      1 MEMYOVKPRHGGAPLRYELPTCMYRLPNVHGSGYGAPEAGHVOEESNLSQALBESROP 60
Db      1 MEMYVKSYGSGAPLRYELPTCMYRLPNVHGSGYGAPEAGHVOEESNLSQALBESROP 60
Qy      61 DILKRLYLKAAVDCGSKMIDTPDADLVNTNIIQADEPTLLTNALDINSVKGKDYALK 120
Db      61 DILKRLYLKAAVDCGSKMIDTPDADLVNTNIIQADEPTLLTNALDINSVKGKDYALK 120

```

```

Db      61 DILKRLYLKAAVDCGSKMIDTPDADLVNTNIIQADEPTLLTNALDINSVKGKDYALK 120
Qy      121 DIVINANPASPPLSLVLRLLCERFVSLVTHSSVKNVPEMLKCGEGQARKQSRHE 180
Db      121 DIVINANPASPPLSLVLRLLCERFVSLVTHSSVKNVPEMLKCGEGQARKQSRHE 180
Qy      181 YOLGFTLLIMKNVPTQMKFSIQTMCPIDEGNIARFLFSLFGQKNAVANATLIDSWDIA 240
Db      181 YOLGFTLLIMKNVPTQMKFSVQTMCPIDEGNIARFLFSLFGQKNAVANATLIDSWDIA 240
Qy      241 IFOLKEGSSKERAAPVRSNMSALGKSPMLAGNELTVADVLMVSVLQQTGGSSGVAAPVAVQ 300
Db      241 MFOLKEGSSKERAAPVRSNMSALGKSPMLAGNELTVADVLMVSVLQQTGGSSGVAAPVAVQ 300
Qy      301 RMWRSCENLAPPTALTKLK 320
Db      301 RMWRSCENLAPPTALTKLK 320

```

```

RESULT 6
Q8R3V2 PRELIMINARY; PRT; 280 AA.
AC 08R3V2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE JTV1.
GN Name=Jtv1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OK NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A..
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Ueffing T.B., Toshvuk S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Richards S., McSwan P.J., McKernan K.U., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,
RA Krzywiński M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A..
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC024480; AAH24480.1; -.
DR WGI: WGI:2385237; JTV1.
DR InterPro: IPR004046; GST_C_1like.
DR InterPro: IPR010987; GST_C_1like.
DR Pfam: PF00043; GST_C_1.
SQ SEQUENCE 280 AA; 31097 MW; B5BE81498963FD2E CRC64;

```

```

Query Match      75.6%; Score 1260; DB 2; Length 280;
Best Local Similarity 88.0%; Pred. No. 5,5e-94;
Matches 243; Conservative 14; Mismatches 19; Indels 0; Gaps 0;

```

```

Qy      45 QESNLSLQALBESQODILKRLYLKAAVDCGSKMIDTPDADLVNTNIIQADEPTLLTN 104

```


Db 5 GETSEPSLQALBSRQDDILKRLYEELKAAVDGSKMHTPDADVDYTNILQADEPTLATN 64
Qy 105 ADDINSVLCGVKALDVIYINNPASPPSLVLAHLLCEHPRVLSVTHHSSVKSPEN 164
Db 65 TLIDINSVLCGVKALDVIYINNPASPPSLVLAHLLCEHPRVLSVTHHSSVKSPEN 124
Qy 165 LKCFGEQKKQKQPRODYOGLFTLIMKNVPTQMKFSIQTMCPLEGSGNIARFLFSLFGOK 224
Db 125 LKCFGEQKKQKQPRODYOGLFTLIMKNVPTQMKFSIQTMCPLEGSGNIARFLFSLFGOK 184
Qy 225 HNAVNATLIDSVWDIAIFQLKEGSSKEKAAVFRSNMSALGKSPWLGNELTVADVVLMSV 284
Db 185 HNAVNATLIDSVWDIAIFQLKEGSSKEKAAVFRSNMSALGKSPWLGNELTVADVVLMSV 244
Qy 285 LQOIGGCSVTVPANVORMRSCENLAPFTALKCLK 320
Db 245 LQOIGGCSVTVPANVORMRSCENLAPFTALKCLK 280

RESULT 7

Q6DR86 PRELIMINARY; PRT; 311 AA.
AC Q6DR86; 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DE 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
GN Name=MG69221; Xenopus tropicalis (western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
NCBI_TaxId=8364;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullen S.J., Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S., Krzywinski M.I., Skalska U., Smallos D.E., Scherch A., Schein J.E., Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RL Klein S., Gerhard D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC074561; AAH74561.1; -
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR010987; GST_C_1like.
DR Pfam; PF00043; GST_C_1.
SQ SEQUENCE 311 AA; 34480 MW; D98F27F73C466154 CRC64;

Query Match 68.7%; Score 1144.5; DB 2; Length 311;
Best Local Similarity 68.8%; Pred. No. 1.5e-84;
Matches 220; Conservative 43; Mismatches 48; Indels 9; Gaps 3;

Qy 1 MEMYQVPEYHGGAPLRYELPTQWRLPNVHGRSYPAGAGHVOESNLSQALBSROD 60
Db 1 MEMYQVPEYHGGAPLRYELPTQWRLPNVHGRSYPAGAGHVOESNLSQALBSROD 52
Qy 61 DILKRLYEELKAAVDGSKMHTPDADVDYTNILQADEPTLATNLDINSVLCGVKAL 120
Db 53 DILKRLYEELKAAVDGSKMHTPDADVDYTNILQADEPTLATNLDINSVLCGVKAL 112
Qy 121 DIVINANPSPPLSLVLAHLLCEHPRVLSVTHHSSVKSPENLKKCFGEQKKQKQPROD 180
Db 113 DIVINANPSPPLSLVLAHLLCEHPRVLSVTHHSSVKSPENLKKCFGEQKKQKQPROD 172
Qy 181 YOLGFTLIMKNVPTQMKFSIQTMCPLEGSGNIARFLFSLFGOKHNAVNATLIDSVWDIA 240
Db 173 YOLGFTLIMKNVPTQMKFSIQTMCPLEGSGNIARFLFSLFGOKHNAVNATLIDSVWDIA 232
Qy 241 ITQLKEGSSKEKAAVFRSNMSALGKSPWLGNELTVADVVLMSVLUOIGGCSVTVPANVQ 300
Db 233 ITQLKEGSSKEKAAVFRSNMSALGKSPWLGNELTVADVVLMSVLUOIGGCSVTVPANVQ 291
Qy 301 RMRSCENLAPFTALKCLK 320
Db 292 RMRSCENLAPFTALKCLK 311

RESULT 8

Q6IND4 PRELIMINARY; PRT; 311 AA.
AC Q6IND4; 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
GN Name=MG68304; Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
NCBI_TaxId=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullen S.J., Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S., Krzywinski M.I., Skalska U., Smallos D.E., Scherch A., Schein J.E., Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RL Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W., Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus Dev. Dym. 225:384-391(2002).
RN [3]

RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RA Klein S., Gerhard D.S.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC072178; AAH72178.1;
 DR InterPro: IPR004046; GST_C1term.
 DR InterPro: IPR010987; GST_C1like.
 DR Pfam: PF00043; GST_C1.
 SQ SEQUENCE 311 AA; 34424 MW; D5E8325C1B88751 CRC64;

Query Match 68.4%; Score 1139.5; DB 2; Length 311;
 Best Local Similarity 69.1%; Pred. No. 3.8e-84;
 Matches 221; Conservative 40; Mismatches 50; Indels 9; Gaps 3;

QY 1 MPYQVRYHGGAPRLVELPTCMYRLPNVHGRSGYPAPAGHVOESNLSIQLESROD 60
 DB 1 MPYQVRYHGGAPRLVELPTCMYRLPNVHGRSGYPAPAGHVOESNLSIQLESROE 52
 QY 61 DILKRIYELKAAVDSKMIQTPDADLVNTIIQADEPTTLTNTALDINSVLGKYGALK 120
 DB 53 DILKRIYELKAAVDSKMIQTPDADLVNTIIQADEPTTLTNTALDINSVLGKYGALK 112
 QY 121 DIVINANPASPPLSLIVLRLLCEHFRVLSYTHSSVSVENILKCGEQNKQPPROD 180
 DB 113 DIVINANPASPPLSLIVLRLLCEHFRVLSYTHSSVSVENILKCGEQNKQPPROE 172
 QY 161 YOLGFTLTKVNPKTQMKFSIQTMCPIDEGGNIAFLFSLFGQKHNANATLIDSWDIA 240
 DB 173 YOLGFTLTKVNPKTQMKFSIQTMCPIDEGGNIAFLFSLFGQKHNANATLIDSWDIA 232
 QY 241 IFQLKESSSKERKAAYFRSNMNSALGKSPMLAGNELTVADVILMSVLOQTGGCSVTYPANVQ 300
 DB 233 IFQLKESSSKERKAAYFRSNMNSALGKSPMLAGNELTVADVILMSVLOQTGGCSVTYPANVQ 291
 QY 301 RWRKSCENLAPNTALIKLK 320
 DB 292 RWRKSCENLAPNTALIKLK 311

RESULT 9

ID 07ZYD7 PRELIMINARY; PRT; 311 AA.
 AC 07ZYD7
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Jv1-prov protein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodinae; Xenopus.
 NC NCB1_TaxID=8355;
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Schenker C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A.C., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M., Skalski U., Smalins D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]

RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius U., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT Initiative";
 RL Dev. Dyn. 225:384-391 (2002).
 RN [3]

RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Strausberg R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC043832; AAH43832.1;
 DR InterPro: IPR004046; GST_C1term.
 DR InterPro: IPR010987; GST_C1like.
 DR Pfam: PF00043; GST_C1.
 SQ SEQUENCE 311 AA; 34647 MW; 2F08C9DA60DD63BE CRC64;

Query Match 67.9%; Score 1132.5; DB 2; Length 311;
 Best Local Similarity 69.1%; Pred. No. 1.4e-83;
 Matches 221; Conservative 39; Mismatches 51; Indels 9; Gaps 4;

QY 1 MPYQVRYHGGAPRLVELPTCMYRLPNVHGRSGYPAPAGHVOESNLSIQLESROD 60
 DB 1 MPYQVRYHGGAPRLVELPTCMYRLPNVHGRSGYPAPAGHVOESNLSIQLESROE 52
 QY 61 DILKRIYELKAAVDSKMIQTPDADLVNTIIQADEPTTLTNTALDINSVLGKYGALK 120
 DB 53 DILKRIYELKAAVDSKMIQTPDADLVNTIIQADEPTTLTNTALDINSVLGKYGALK 112
 QY 121 DIVINANPASPPLSLIVLRLLCEHFRVLSYTHSSVSVENILKCGEQNKQPPROD 180
 DB 113 DIVINANPASPPLSLIVLRLLCEHFRVLSYTHSSVSVENILKCGEQNKQPPROE 172
 QY 161 YOLGFTLTKVNPKTQMKFSIQTMCPIDEGGNIAFLFSLFGQKHNANATLIDSWDIA 240
 DB 173 YOLGFTLTKVNPKTQMKFSIQTMCPIDEGGNIAFLFSLFGQKHNANATLIDSWDIA 232
 QY 241 IFQLKESSSKERKAAYFRSNMNSALGKSPMLAGNELTVADVILMSVLOQTGGCSVTYPANVQ 300
 DB 233 IFQLKESSSKERKAAYFRSNMNSALGKSPMLAGNELTVADVILMSVLOQTGGCSVTYPANVQ 291
 QY 301 RWRKSCENLAPNTALIKLK 320
 DB 292 RWRKSCENLAPNTALIKLK 311

RESULT 10

ID 07T3C0 PRELIMINARY; PRT; 321 AA.
 AC 07T3C0
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Zgc:63976.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NC NCB1_TaxID=7955;
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Schenker C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

| Query Match | Similarity | 53.2% | Score 887.5 | DB 2 | Length 321 |
|-------------|------------|--------------|-------------------|------------|------------|
| Best Local | Similarity | 55.6% | Pred. No. 1,1e-63 | | |
| Matches | 184 | Conservative | 50 | Mismatches | 76 |
| | | | | Indels | 21 |
| | | | | Gaps | 8 |

| Query | 1 | MPMYQVPRYHGCGAPLRYELPTCYRLRPNVHGSRGSPAPAGHVOE-ESNLSLOALBESRQ | 59 |
|-------|-----|--|-----|
| Db | 1 | MPMYQVPRV--SPADITVDLPTCMWKLPNVHAQ--GASLGERHALNGEVDPEYVALERQ | 56 |
| Qy | 60 | DDIKRLRYELKAAVDGSKMLQTPPADLDVNIIGADEFTLT-----TNMLDINSV | 111 |
| Db | 57 | DELRKLYELKAAVYDGLAKVTITPADDASTLAH----TLTTPPADAVLRGNADLDD | 111 |
| Qy | 112 | LGXQYGLKDIVINANPASPPLSLVLRLLCEHRYLSTVHTHSYKSVENILKCPGE | 171 |
| Db | 112 | LXKOLGALRDIIVANPAQPSLILVLRLLCQRFQVLSVHVSISVETVPAPLLSCIGP | 171 |
| Qy | 172 | QN-KKQPRQDVQLGFTLLIKRVPTQMKFSIQTMCIPEGENINARFLFSLG-QKANVN | 229 |
| Db | 172 | RHTSYVAHRPQLGTLTIWKDVSKIQMKFSIQNMCPLEGENVARFLYRLLAGAPRDRVS | 231 |
| Qy | 230 | ATLIDSWDIALPOLKEGSSKEKAAPFRSNMSALGKSPMLGNELTVADVLTWLSYLAQIG | 289 |
| Db | 232 | ATLMDGWVDLTFQALGEGSKERAAYLRALNALGRSPWLLGQFESLADIYSACCVIQTG | 291 |
| Qy | 290 | GCSYTVPRANVQRMKRSCENLAPFNTALKLK | 320 |
| Db | 292 | QTS-SAPANVQRMKRSCENLGYFSCVDPBLQ | 321 |

| RESULT 11 | ID | 0707A3 | PRELIMINARY | PRT | 340 AA. |
|-----------|--|---|-------------|-----|---------|
| AC | 0707A3 | | | | |
| DT | 01-MAR-2004 | (TREMBLrel. 26, Created) | | | |
| DT | 01-MAR-2004 | (TREMBLrel. 26, last sequence update) | | | |
| DT | 01-MAR-2004 | (TREMBLrel. 26, last annotation update) | | | |
| DE | AGCP5808 | (Fragment) | | | |
| GN | Name=agCG50514; ORFName=ENSAAG00000011827; | | | | |
| OS | Anopheles gambiae str. PEST. | | | | |
| OC | Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preygota; | | | | |
| NC | Nooptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anophelae. | | | | |
| OX | NCBI_TaxID=180454; | | | | |
| NP | [1] | | | | |
| | SEQUENCE FROM N.A. | | | | |

| | | | | |
|-----------------------|------------------|-------------------|------------|-------------|
| Query Match | 16.0% | Score 267.5; | DB 2; | Length 340; |
| Best Local Similarity | 29.2% | Pred. No.2.6e-13; | | |
| Matches 100; | Conservative 47; | Mismatches 133; | Indels 63; | Gaps 13; |

| | |
|-----------|--------------|
| RESULT 12 | |
| Q7KUM5 | |
| ID Q7KUM5 | PRELIMINARY; |
| | PRT; 301 AA. |

DT 05-JUL-2004 (Tremblay, 27, Created)
DT 05-JUL-2004 (Tremblay, 27, Last sequence update)
DT 05-JUL-2004 (Tremblay, 27, Last annotation update)
DE CG12304-PB.
GN
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Suton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blatz R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gaber G.L.,
RA Abil J.F., Aghayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhargava D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burris K.C., Busam D.A., Butler H., Cadien J.E., Center A., Chandra I.,
RA Cherry J.M., Chesnut S., Dahle C., Davenport L.B., Davies P.,
RA de Paolis B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match | 13.4% | Score 224; | DB 2; | Length 301; |
| Best Local Similarity | 28.1%; | Pred. No. 7.6e-10; | | |
| Matches 85; | Conservative 47; | Mismatches 120; | Indels 50; | Gaps 1 |

| Query Match | 13.4% | Score 224 | DB 2 | Length 313 |
|---|-------|-----------------|------|-------------------------------------|
| Best Local Similarity | 28.1% | Pred. No. 8e-10 | | |
| Matches | 85 | Conservative | 47 | Mismatches 120; Indels 50; Gaps 10; |
| 18 VELPTCMYRLPNV-----HGRSYGAPAPAGVQVEESNLISQLESRODITLKELYEI 69 | ::: | ::: | ::: | ::: |
| 25 IKLPQWPIKKNVSLAIDSLASGSSSTASSTASSTSCDDPAVA---RQEKTKOLEEL 81 | ::: | ::: | ::: | ::: |
| 70 KAANDGLSKMTQTPDADIDVTNIIIQADEPTLTITNALDNLVSGKDYGALKEIVIANEFA 129 | ::: | ::: | ::: | ::: |
| 82 KQIQGLQIR-----AGIGVCG--KTPQHTTARQNG-----GLKEVPLQDVGIVNGHEN 125 | ::: | ::: | ::: | ::: |
| 130 SEPPLSLVIRHLCEHFRLVSTVHTHSVSYSV-----PENILKCFGQNKQKPPRODYQ 182 | ::: | ::: | ::: | ::: |
| 126 FIPVALLATLAKRAMNNLTIDVKTFTTSHVIAADIDAPAAEFENAIATK--EVNPAALPK--- 179 | ::: | ::: | ::: | ::: |

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:25:39 ; Search time 70.371 Seconds
(without alignments)
1758.725 Million cell updates/sec

Title: US-10-622-817-5

Perfect score: 1655
Sequence: 1 MPYGVKPRHGSAPLRLVEL.....RLKSGENLAFETALQLLK 320

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query | Length | DB | ID | Description |
|------------|-------|-------|--------|----|----------|---------------------|
| 1 | 1464 | 88.5 | 320 | 7 | AD58613 | Ad58613 Human Pro |
| 2 | 1464 | 88.5 | 320 | 8 | ABM60843 | Abm60843 Tumour-as |
| 3 | 1436 | 86.8 | 312 | 2 | AAW25776 | AAW25776 JTV1 Prot |
| 4 | 1436 | 86.8 | 312 | 8 | ADR86551 | ADR86551 1-312 am |
| 5 | 1232 | 74.4 | 272 | 7 | ADCI0204 | ADCI0204 Human NOV |
| 6 | 1069 | 64.6 | 229 | 8 | ADR86551 | ADR86551 84-312 am |
| 7 | 735 | 44.4 | 161 | 8 | ADR86552 | ADR86552 1-161 am |
| 8 | 248 | 15.0 | 51 | 5 | ABP01502 | ABP01502 Human CRF |
| 9 | 216.5 | 13.1 | 334 | 4 | AB862468 | AB862468 Drosophila |
| 10 | 148 | 8.9 | 263 | 4 | ABG23964 | ABG23964 Novel hum |
| 11 | 126.5 | 7.6 | 716 | 3 | AA641411 | AA641411 Arabidops |
| 12 | 126.5 | 7.6 | 719 | 3 | AA641410 | AA641410 Arabidops |
| 13 | 126.5 | 7.6 | 748 | 3 | AA641409 | AA641409 Arabidops |
| 14 | 106 | 6.4 | 1512 | 4 | AAU04349 | AAU04349 Mammalian |
| 15 | 106 | 6.4 | 1512 | 4 | AAW78732 | AAW78732 Human pro |
| 16 | 106 | 6.4 | 1550 | 4 | AAW79716 | AAW79716 Human pro |
| 17 | 104 | 6.2 | 980 | 8 | ADN23683 | ADN23683 Bacterial |
| 18 | 102 | 6.3 | 1049 | 6 | ABM67090 | ABM67090 Phototrab |
| 19 | 101.5 | 6.1 | 724 | 4 | AA685046 | AA685046 Strimp wh |
| 20 | 99 | 6.0 | 691 | 6 | ABU16445 | ABU16445 Protein e |
| 21 | 98 | 5.9 | 243 | 4 | AAU58158 | AAU58158 Propionib |
| 22 | 98 | 5.9 | 243 | 4 | ABM54677 | ABM54677 Propionib |
| 23 | 98 | 5.9 | 644 | 6 | ABJ18952 | ABJ18952 Pathogen |
| 24 | 98 | 5.9 | 690 | 6 | AAU34345 | AAU34345 Stephyloc |
| 25 | 98 | 5.9 | 690 | 6 | ABM71749 | ABM71749 Stephyloc |

| | | | | | | |
|----|------|-----|------|---|----------|--------------------|
| 26 | 98 | 5.9 | 985 | 8 | ADN88622 | Adn88622 Rat epid |
| 27 | 98 | 5.9 | 997 | 8 | ADN88620 | Adn88620 Rat epid |
| 28 | 98 | 5.9 | 1004 | 8 | ADN88621 | Adn88621 Rat epid |
| 29 | 98 | 5.9 | 1013 | 8 | ADN88594 | Adn88594 Rat epid |
| 30 | 98 | 5.9 | 1016 | 8 | ADN88619 | Adn88619 Rat epid |
| 31 | 98 | 5.9 | 1032 | 8 | ADN88618 | Adn88618 Rat epid |
| 32 | 96.5 | 5.8 | 1398 | 8 | ABM83058 | Abm83058 Human dia |
| 33 | 96.5 | 5.8 | 1398 | 8 | ABM83057 | Abm83057 Human dia |
| 34 | 96.5 | 5.8 | 1440 | 8 | ADN03674 | Adn03674 Antipeori |
| 35 | 96.5 | 5.8 | 1440 | 8 | ADP54124 | Adp54124 Human pro |
| 36 | 96.5 | 5.8 | 1440 | 8 | ADP23041 | Adp23041 PRO polyP |
| 37 | 96 | 5.8 | 2273 | 2 | AA898811 | AA898811 Eryalphe |
| 38 | 95 | 5.7 | 722 | 8 | AD820980 | AD820980 Bacterial |
| 39 | 94 | 5.7 | 436 | 5 | AB854068 | AB854068 Lactococc |
| 40 | 93.5 | 5.6 | 1049 | 5 | AB891239 | AB891239 Herbicida |
| 41 | 93.5 | 5.6 | 1064 | 5 | AAE25162 | AAE25162 RCH1.5 pr |
| 42 | 93.5 | 5.6 | 1090 | 4 | ABG17694 | ABg17694 Novel hum |
| 43 | 93.5 | 5.6 | 1090 | 4 | ABG18331 | ABg18331 Novel hum |
| 44 | 93 | 5.6 | 204 | 8 | AD821772 | AD821772 Bacterial |
| 45 | 93 | 5.6 | 559 | 2 | AA810682 | AA810682 Polymydr |

ALIGNMENTS

RESULT 1
AD58613
ID AD58613 standard; protein; 320 AA.
XX
AC AD58613;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein Q13155, SEQ ID NO 4489.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX
PA (FARB) BAYER AG.
XX
PI Woolf C, D'ureo D, Befort K, Coeligan M;
XX
DR WPI, 2003-268312/26.
XX
DR GENBANK; Q13155.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1, Page: 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal

CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SN1)) and the compound that
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPD at
CC ftp.wipo.int/pub/published_pct_sequences.

| | |
|--|--|
| ID | AAM25776 standard; protein; 312 AA. |
| XX | |
| AC | AAM25776; |
| XX | |
| DT | 19-DEC-1997 (first entry) |
| XX | |
| DE | JTV1 protein. |
| XX | |
| KW | JTV1; hPMS2; probe; detection; chromosome 7; deletion; mismatch repair gene; hereditary non-polyposis colorectal cancer; homologous recombination. |
| KM | |
| XX | |
| OS | Homo sapiens. |
| XX | |
| PN | MO9708312-A1. |
| PD | 06-MAR-1997. |
| XX | |
| PE | 26-AUG-1996; 96WO-US013598. |
| PR | 24-AUG-1995; 95US-00518862. |
| XX | |
| PA | (UYJO) UNIV JOHNS HOPKINS. |
| PI | Vogelstein B, Kinzler KW, Nicolaides NC; WPI; 1997-179269/16. N-PsDB; AAT86182. |
| DR | |
| XX | |
| PT | Novel chromosome 7 gene, JTV1 - used for detecting chromosome 7 deletions, and PMS2 promoter activity. |
| FT | |
| XX | |
| PS | Claim 5; Fig 2; 31pp; English. |
| XX | |
| CC | This sequence is JTV1 protein and is encoded by DNA isolated from human chromosome 7. The JTV1 coding sequence is located upstream from hPMS2. JTV1 cDNA can be used as probes to detect chromosome 7 deletions involving JTV1. Due to the overlapping promoter regions, deletions of JTV1 would also affect PMS2 (a mismatch repair gene) expression, leading to hereditary non-polyposis colorectal cancer. JTV1 can also be used to assay activity or competence of the PMS2 promoter region, the presence of JTV1 suggesting that the PMS2 promoter is intact. JTV1 sequences can also be used to guide homologous recombination at the PMS2 locus |
| CC | |
| XX | |
| SQ | Sequence 312 AA; |
| Query Match | 86.8%; Score 1436; DB 2; Length 312; |
| Match Local Similarity | 87.5%; Pred. No. 5,6e-139; |
| Matches 273; Conservative 14; Mismatches 25; Indels 0; Gaps 0; | |
| OY | 1 MPMYVVKRPHGGSAFLRYELPTCMYRLPNVHSKTSPATDAGHVQETSEPSIQALESROD 60 |
| Db | 1 MPMYVVKRPHGGGALRYELPTCMYRLPNVHGRSGTPAGCHVQEESNLISQALESRÖD 60 |
| OY | 61 DILKRLYEIKAAVDGLSKMIHTPDADLDVTNILOADEPTTLATNTLADNSVLGKDYGALK 120 |
| Db | 61 DILKRLYEIKAAVDGLSKMIQTDPADLDVTNILOADEPTTLTNNLADNSVLGKDYGALK 120 |
| OY | 121 DIVINANPASPLSLVLVHRLLCEFRVYLSTVHTSSVKSVPENLLKCFGEONKKQPROD 180 |
| Db | 121 DIVINANPASPLSLVLVHRLLCEFRVYLSTVHTSSVKSVPENLLKCFGEONKKQPROD 180 |
| OY | 181 YOLGFTLLMKVNPKTOMKPSVOTMCPRIEGEIGNIAFLFSLFGQKNNAATLTLLIDSWVIA 240 |
| Db | 181 YOLGFTLLMKVNPKTOMKPSIOTMCPRIEGEIGNIAFLFSLFGQKNNAATLTLLIDSWVIA 240 |
| OY | 241 MFQLREGSSKEKAVERSNASLAGRSFMLVGNEILTAVADVLMVSVLQQTGGSSGAATTNVQ 300 |
| Db | 241 IFQLREGSSKEKAVERSNASLAGRSFMLVGNELTVADVLMVSVLQQTGGSVTPAVNVQ 300 |
| OY | 301 RWLKSCENLAAP 312 |
| Db | 301 RWRMSCENLAAP 312 |

```

RESULT 4
ADNR86551
ID ADNR86551 standard; protein; 312 AA.
XX AC ADNR86551;
XX DT 18-NOV-2004 (first entry)
XX DE 1-312 amino acid sequence of p38/JTV-1 protein.
XX KM p38/JTV-1; Cytoebatic; cancer; leukemia; anticancer.
XX OS Homo sapiens.
XX PN EP1454628-A2.
XX PD 08-SEP-2004.
XX PF 09-SEP-2003; 2003EP-00020344.
XX PR 03-MAR-2003; 2003XR-00013058.
XX PA (UYSE-) UNIV SEOUL NAT IND FOUND.
XX PI Kim S, Park B;
XX DR MPI; 2004-627822/61.
XX DR N-PESD; ADNR86548.
PT New isolated p38/JTV-1 protein, useful as medicament for treating cancer
PT e.g., stomach, liver, blood, bone, pancreatic, skin, head or neck cancer
PT and cutaneous or intraocular melanoma, as well as for screening new
PT anticancer agents.
PS Claim 5; SEQ ID NO 4; 47pp; English.
XX
XX The present invention relates to an isolated p38/JTV-1 protein for use as
CC medicament. The p38/JTV-1 protein or the pharmaceutical composition is
CC useful as medicament for treating breast cancer, large intestinal cancer,
CC lung cancer, small cell lung cancer, stomach cancer, liver cancer, blood
CC cancer, bone cancer, pancreatic cancer, skin cancer, head or neck cancer,
CC cutaneous or intraocular melanoma, uterine sarcoma, ovarian cancer,
CC rectal cancer, anal cancer, colon cancer, fallopian tube carcinoma,
CC endometrial carcinoma, cervical cancer, vulval cancer, vaginal carcinoma,
CC Hodgkin's disease, esophageal cancer, small intestine cancer, endocrine
CC cancer, thyroid cancer, parathyroid cancer, adrenal cancer, soft tissue
CC tumour, uethelial cancer, penile cancer, prostate cancer, chronic or acute
CC leukemia, lymphocytic lymphoma, bladder cancer, kidney cancer, ureter
CC cancer, renal cell carcinoma, renal pelvic carcinoma, CNS tumour, primary
CC CNS lymphoma, bone marrow tumour, brain stem nerve gliomas, pituitary
CC adenoma, or their combination. The protein is useful as a target for
CC screening new anticancer agents. The present sequence represents the 1-
CC 312 amino acid sequence of p38/JTV-1 protein.
XX
XX
SEQ Sequence 312 AA;
Query Match 86.8%; Score 1436; DB 8; Length 312;
Best Local Similarity 87.5%; Pred. No. 5,66-119;
Matches 273; Conservative 14; Mismatches 25; Indels 0; Gaps 0;
QY 1 MPMTQVRYHGSGAPLRYELPTCMYRLPNVHSGKTSPTATAGHVOESEPGLAESROD 60
DB 1 MPMTQVRYHGSGAPLRYELPTCMYRLPNVHSGRYGPAQGHVOESENUSLOLESROD 60
QY 61 DILKRLVELKAADVGLSKMTHTPPADLDVNTIIQADSEPTTLATNTLDLSNYGKDYGALK 120
DB 61 DILKRLVELKAADVGLSKMTHTPPADLDVNTIIQADSEPTTLATNTLDLSNYGKDYGALK 120
QY 121 DIVINANPASPPLSLVTLHRLCGRYVLSVYTHSSVYKVPENLVYCFGQAKKSRHE 180
DB 121 DIVINANPASPPLSLVTLHRLCGRYVLSVYTHSSVYKVPENLVYCFGQAKKSRHE 180
QY 121 DIVINANPASPPLSLVTLHRLCGRYVLSVYTHSSVYKVPENLVYCFGQAKKSRHE 180
DB 121 DIVINANPASPPLSLVTLHRLCGRYVLSVYTHSSVYKVPENLVYCFGQAKKSRHE 180

```

```
QY 181 YQGLTILMKNVKPTQMKTSVOTMCPTEGEGNIARFLFSLFGOKNAVTLTIDSWDIA 240
XX |||||
XX 181 YQGLTILMKNVKPTQMKTSVOTMCPTEGEGNIARFLFSLFGOKNAVTLTIDSWDIA 240
Db 181 YQGLTILMKNVKPTQMKTSVOTMCPTEGEGNIARFLFSLFGOKNAVTLTIDSWDIA 240
QY 241 MFLRGGSSKEKAARFRSMNSALGRSPMLVGNELTVADVVLMSVLCQRTGSSGAAPTAVQ 300
Db 241 IFQLKGGSSKEKAARFRSMNSALGRSPMLVGNELTVADVVLMSVLCQRTGSSGAAPTAVQ 300
QY 301 RMLKSCENLAPF 312
Db 301 RMLKSCENLAPF 312
Db 301 RMLKSCENLAPF 312
RESULT 5
ADCI0204
ID ADCI0204 standard; protein; 272 AA.
AC ADCI0204;
XX
XX
XX 18-DEC-2003 (first entry)
DT
XX
XX Human NOVX polypeptide SEQ ID NO: 226.
DE
XX
XX cytosolic; antidiabetic; anorectic; cerebroprotective; neuroprotective;
KW antiinflammatory; gene therapy; antisense therapy; thymimetic; NOVX;
KW pathology; cancer; diabetes; obesity; endocrine disorder; CNS disorder;
KW inflammatory disorder; chromosome mapping; tissue typing;
KW predictive medicine.
XX
XX Homo sapiens.
OS
XX
XX WO200300842-A2.
XX
XX 03-JAN-2003.
XX
XX
XX 04-JUN-2002; 2002WO-US017443.
PF
XX
XX 04-JUN-2001; 2001US-0295607P.
PR 04-JUN-2001; 2001US-0295607P.
PR 06-JUN-2001; 2001US-0296404P.
PR 06-JUN-2001; 2001US-0296404P.
PR 07-JUN-2001; 2001US-0296575P.
PR 11-JUN-2001; 2001US-0297414P.
PR 12-JUN-2001; 2001US-0295573P.
PR 12-JUN-2001; 2001US-0297567P.
PR 14-JUN-2001; 2001US-0298285P.
PR 15-JUN-2001; 2001US-0298285P.
PR 18-JUN-2001; 2001US-0299133P.
PR 19-JUN-2001; 2001US-0299230P.
PR 21-JUN-2001; 2001US-0299499P.
PR 22-JUN-2001; 2001US-0300177P.
PR 26-JUN-2001; 2001US-0300883P.
PR 28-JUN-2001; 2001US-0301530P.
PR 28-JUN-2001; 2001US-0301530P.
PR 03-JUL-2001; 2001US-0302951P.
PR 31-JUL-2001; 2001US-0308890P.
PR 14-SEP-2001; 2001US-0322297P.
PR 25-SEP-2001; 2001US-0324699P.
PR 03-DEC-2001; 2001US-0337477P.
PR 14-DEC-2001; 2001US-0341562P.
PR 21-FEB-2002; 2002US-0358566P.
PR 21-FEB-2002; 2002US-0358566P.
PR 22-FEB-2002; 2002US-0359122P.
PR 22-FEB-2002; 2002US-0358978P.
PR 22-FEB-2002; 2002US-0359034P.
PR 22-FEB-2002; 2002US-0359121P.
PR 27-FEB-2002; 2002US-0359964P.
PR 01-MAR-2002; 2002US-0360858P.
PR 12-MAR-2002; 2002US-0363430P.
PR 12-MAR-2002; 2002US-0363676P.
PR 10-APR-2002; 2002US-0371346P.
PR 10-MAY-2002; 2002US-0379444P.
XX
```

```
PR 04-JUN-2002; 2002US-00379444.
XX
XX (CURA-) CURAGEN CORP.
XX
XX
PI Agee ML, Anderson DM, Berghs C, Caeman SJ, Catterton E;
PI Delpo VA, Edinger SR, Eisen A, Ellemann K, Gangolli EA;
PI Gerlach VL, Gorman L, Guo X, Herrmann JL, Hjalte T, Ji W, Kekuda R;
PI Khrantsov NV, Li L, Liu X, Malyanar UM, Miller CE, Miller I;
PI Ott T, Padigaru M, Patturajan M, Pena CB, Rastelli L, Rieger DK;
PI Rothenberg ME, Shenoy SG, Shimkets RA, Smithson G, Spaderna SK;
PI Splet KA, Stone DJ, Vernet CM, Zhong H, Zhong W, Alsbrook JP;
PI Burgess CE, Lepley DM;
XX
DR WPI, 2003-210149/20.
DR N-PSDB; ADCI0203.
XX
XX New isolated NOVX polypeptides and nucleic acid molecules useful for
PT treating, preventing and diagnosing pathological conditions with NOVX-
PT associated disorders, such as cancer, obesity, diabetes and inflammatory
PT or CNS diseases.
XX
XX Claim 1; SEQ ID NO 226; 772bp; English.
XX
XX
XX The invention relates to novel isolated polypeptides, mature form of the
CC polypeptide, a sequence that is 95% identical to the polypeptide or the
CC polypeptide comprising one or more conservative substitutions. The NOVX
CC polypeptide is useful for treating or preventing a pathology associated
CC with the polypeptide e.g. disorders associated with aberrant expression
CC or activity of the polypeptide, such as cancer, diabetes, obesity, and
CC endocrine, CNS and inflammatory disorders. They can also be used in
CC various detection and screening assays, chromosome mapping, tissue typing
CC and predictive medicine. This sequence corresponds to one of the
CC polypeptides of the invention.
XX
XX Sequence 272 AA;
SQ
Query Match 74.4%; Score 1232; DB 7; Length 272;
Best Local Similarity 88.4%; Pred. No 5,1e-118;
Matches 237; Conservative 12; Mismatches 19; Indels 0; Gaps 0;
QY 45 QETSEPSIQLESFODDILKRYELKAAVDGLSKMTHTPDADVDVNTIQADEPTLATN 104
Db 5 QESNLSIQLESFODDILKRYELKAAVDGLSKMTHTPDADVDVNTIQADEPTLATN 64
QY 105 TLIDNSVYGKDYGLKQIVINANPASPPLSLVTHLLCERYVSLVTVTHSSVKNVPPN 164
Db 65 ALDINSVYGKDYGLKQIVINANPASPPLSLVTHLLCERYVSLVTVTHSSVKNVPPN 124
QY 165 LVKCFGQARKQSRHEVYGLFTLIMKNVPTQMKTSVOTMCPTEGEGNIARFLFSLFGOK 224
Db 125 LVKCFGQARKQSRHEVYGLFTLIMKNVPTQMKTSVOTMCPTEGEGNIARFLFSLFGOK 184
QY 225 HNAVTLTIDSWDIAFOLRFGSSKEKAARFRSMNSALGRSPMLVGNELTVADVVLMSV 284
Db 185 HNAVTLTIDSWDIAFOLRFGSSKEKAARFRSMNSALGRSPMLVGNELTVADVVLMSV 244
QY 285 LQQTGSSGAAPTAVQRLKSCENLAPF 312
Db 245 LQQTGSSGAAPTAVQRLKSCENLAPF 272
RESULT 6
ADCI0204
ID ADCI0204 standard; protein; 272 AA.
AC ADCI0204;
XX
XX 18-NOV-2004 (first entry)
DT
XX
XX 84-312 amino acid sequence of p38/JTV-1 protein.
DE
XX
XX p38/JTV-1; Cytostatic; cancer; leukemia; anticancer.
XX
```

OS Homo sapiens.
 XX
 PN EP1454628-A2.
 XX
 PD 08-SEP-2004.
 XX
 PF 09-SEP-2003; 2003EP-00020344.
 XX
 PR 03-MAR-2003; 2003KR-00013058.
 XX
 PA (UYSE-) UNIV SEOUL NAT IND FOUND.
 XX
 PI Kim S, Park B;
 XX
 DR MPI; 2004-627822/61.
 DR N-PSDB; ADR86550.
 XX
 PT New isolated p38/JTV-1 protein, useful as medicament for treating cancer
 PT e.g., stomach, liver, blood, bone, pancreatic, skin, head or neck cancer
 PT and cutaneous or intraocular melanoma, as well as for screening new
 PT anticancer agents.
 XX
 PS Claim 5; SEQ ID NO 6; 47bp; English.
 XX
 CC The present invention relates to an isolated p38/JTV-1 protein for use as
 CC medicament. The p38/JTV-1 protein or the pharmaceutical composition is
 CC useful as medicament for treating breast cancer, large intestinal cancer,
 CC lung cancer, small cell lung cancer, stomach cancer, liver cancer, blood
 CC cancer, bone cancer, pancreatic cancer, skin cancer, head or neck cancer,
 CC cutaneous or intraocular melanoma, uterine sarcoma, ovarian cancer,
 CC rectal cancer, anal cancer, colon cancer, fallopian tube carcinoma,
 CC endometrial carcinoma, cervical cancer, vulval cancer, vaginal carcinoma,
 CC Hodgkin's disease, esophageal cancer, small intestine cancer, endocrine
 CC cancer, thyroid cancer, parathyroid cancer, adrenal cancer, soft tissue
 CC tumour, urethral cancer, penile cancer, prostate cancer, chronic or acute
 CC leukemia, lymphocytic lymphoma, bladder cancer, kidney cancer, ureter
 CC cancer, renal cell carcinoma, renal pelvic carcinoma, CNS tumour, primary
 CC CNS lymphoma, bone marrow tumour, brain stem nerve glioma, pituitary
 CC adenoma, or their combination. The protein is useful as a target for
 CC screening new anticancer agents. The present sequence represents the 84-
 CC 312 amino acid sequence of p38/JTV-1 protein.
 CC
 XX
 SQ Sequence 229 AA;
 Query Match 64.6%; Score 1069; DB 8; Length 229;
 Best Local Similarity 88.2%; Pred. No. 2.6e-101;
 Matches 202; Conservative 12; Mismatches 15; Indels 0; Gaps 0;
 QY 84 DADLDVTNTLQADEPTTLATNTLDLNSVVGKDYGALKDIVINANPASPPSLVLRLLC 143
 DB 1 DADLDVTNTLQADEPTTLATNTLDLNSVVGKDYGALKDIVINANPASPPSLVLRLLC 60
 QY 144 EERYRLSTVHTSHSVKNVENVKCFGEQARKQSREHYQLGFTLIWKVNPKTOMKESVOT 203
 DB 61 EHFRLVSTVHTSHSVKSVPENLKCFGEQARKQKQRPDYQLGFTLIWKVNPKTOMKESVOT 120
 QY 204 MCPPIEGEGNIARPLPSLFGQKNAVNTLTIDSVVDIAMFQLRGSGSKKGAVFRSNLSL 263
 DB 121 MCPPIEGEGNIARPLPSLFGQKNAVNTLTIDSVVDIAMFQLRGSGSKKGAVFRSNLSL 180
 QY 264 GRSFPLVGNELTVADVLTMSVLTQQTGSSGAAPNTNQRWLKSCENIAPF 312
 DB 181 GRSFPLVGNELTVADVLTMSVLTQQTGSSGAAPNTNQRWLKSCENIAPF 229

DE 1-161 amino acid sequence of p38/JTV-1 protein.
 XX
 KM p38/JTV-1; Cytostatic; cancer; leukemia; anticancer.
 XX
 OS Homo sapiens.
 XX
 PN EP1454628-A2.
 XX
 PD 08-SEP-2004.
 XX
 PF 09-SEP-2003; 2003EP-00020344.
 XX
 PR 03-MAR-2003; 2003KR-00013058.
 XX
 PA (UYSE-) UNIV SEOUL NAT IND FOUND.
 XX
 PI Kim S, Park B;
 XX
 DR MPI; 2004-627822/61.
 DR N-PSDB; ADR86549.
 XX
 PT New isolated p38/JTV-1 protein, useful as medicament for treating cancer
 PT e.g., stomach, liver, blood, bone, pancreatic, skin, head or neck cancer
 PT and cutaneous or intraocular melanoma, as well as for screening new
 PT anticancer agents.
 XX
 PS Claim 5; SEQ ID NO 5; 47bp; English.
 XX
 CC The present invention relates to an isolated p38/JTV-1 protein for use as
 CC medicament. The p38/JTV-1 protein or the pharmaceutical composition is
 CC useful as medicament for treating breast cancer, large intestinal cancer,
 CC lung cancer, small cell lung cancer, stomach cancer, liver cancer, blood
 CC cancer, bone cancer, pancreatic cancer, skin cancer, head or neck cancer,
 CC cutaneous or intraocular melanoma, uterine sarcoma, ovarian cancer,
 CC rectal cancer, anal cancer, colon cancer, fallopian tube carcinoma,
 CC endometrial carcinoma, cervical cancer, vulval cancer, vaginal carcinoma,
 CC Hodgkin's disease, esophageal cancer, small intestine cancer, endocrine
 CC cancer, thyroid cancer, parathyroid cancer, adrenal cancer, soft tissue
 CC tumour, urethral cancer, penile cancer, prostate cancer, chronic or acute
 CC leukemia, lymphocytic lymphoma, bladder cancer, kidney cancer, ureter
 CC cancer, renal cell carcinoma, renal pelvic carcinoma, CNS tumour, primary
 CC CNS lymphoma, bone marrow tumour, brain stem nerve glioma, pituitary
 CC adenoma, or their combination. The protein is useful as a target for
 CC screening new anticancer agents. The present sequence represents the 1-
 CC 161 amino acid sequence of p38/JTV-1 protein.
 CC
 XX
 SQ Sequence 161 AA;
 Query Match 44.4%; Score 735; DB 8; Length 161;
 Best Local Similarity 88.8%; Pred. No. 4.4e-67;
 Matches 143; Conservative 5; Mismatches 13; Indels 0; Gaps 0;
 QY 1 MEMYQVPRYHGSGAPLRVELPTCMYRLPNVHSKTTSPATDAGHGETSPPSLQALSSROD 60
 DB 1 MEMYQVPRYHGSGAPLRVELPTCMYRLPNVHSKTTSPATDAGHGETSPPSLQALSSROD 60
 QY 61 DILKRLYEKAAVNDGSKMHTPPDADLDVTNTLQADEPTTLATNTLDLNSVVGKDYGALK 120
 DB 61 DILKRLYEKAAVNDGSKMHTPPDADLDVTNTLQADEPTTLATNTLDLNSVVGKDYGALK 120
 QY 121 DIVINANPASPPSLVLRLLCERVRVSTVHTSHSVKNV 161
 DB 121 DIVINANPASPPSLVLRLLCERVRVSTVHTSHSVKNV 161

RESULT 7
 ADR86552
 ID ADR86552 standard; protein; 161 AA.
 XX
 AC ADR86552;
 XX
 DT 18-NOV-2004 (first entry)
 XX

RESULT 8
 ABP01502
 ID ABP01502 standard; protein; 51 AA.
 XX
 AC ABP01502;
 XX
 DT 24-JUN-2002 (first entry)
 XX

DE Human ORFX protein sequence SEQ ID NO:2986.
XX
KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis.
XX
OS Homo sapiens.
XX MO200192523-A2.
XX
XX 06-DEC-2001.
XX
XX 29-MAY-2001; 2001WO-US010836.
XX
XX 30-MAY-2000; 2000US-0206132P.
XX 29-AUG-2000; 2000US-0228716P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shinkets RA, Leach MD;
XX
XX WPI; 2002-106308/14.
XX N-PSDB; ABN17254.
XX
XX Novel human polypeptides and polynucleotides useful for diagnosing,
XX preventing and treating cardiovascular disease, neurodegenerative,
XX hyperproliferative disorders and autoimmune disorders.
XX
XX Disclosure; SEQ ID NO 2986; 1037bp; English.
XX
XX The present invention describes substantially purified human proteins
XX (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
XX in the specification). ABN15762 to ABN27252 encode the human ORFX.
XX CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
XX treating or preventing a pathology associated with an ORFX-associated
XX disorder in humans, and in the manufacture of a medicament for treating a
XX syndrome associated with ORFX-associated disorder. ORFX polynucleotide
XX sequences can be used in gene therapy. ORFX sequences can be used in the
XX treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
XX psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
XX osteoarthritis, neurodegenerative disorders, disorders related to organ
XX transplantation, cardiovascular diseases, diabetes mellitus, systemic
XX lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
XX storage disease, various immune deficiencies and disorders, infectious
XX diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
XX arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
XX disease and autoimmune inflammatory eye disease. ORFX proteins are also
XX useful for treating burns, incisions, ulcers, for treating osteoporosis,
XX bone degenerative disorders, or periodontal disease, and for gut
XX protection or regeneration and treatment of lung or liver fibrosis,
XX reperfusion injury in various tissues and conditions resulting from
XX systemic cytokine damage. N.B. The sequence data for this patent did not
XX form part of the printed specification, but was obtained in electronic
XX format directly from WIPO at ftp.wipo.int/pub/published_sequences
SQ Sequence 51 AA;
Query Match 15.0%; Score 248; DB 5; Length 51;
Best Local Similarity 96.1%; Pred. No. 1.3e-17;
Matches 49; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 74 DGLSKMIRPDADLDVTNIIQADEPTTLATNTLINSVLGKDYGLAKDIOVI 124
DB 1 DGLSKMIRPDADLDVTNIIQADEPTTLATNTLINSVLGKDYGLAKDIOVI 51
RESULT 9
ABB62468

ID ABB62468 standard; protein; 334 AA.
XX
XX AC ABB62468;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 14196.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
XX Drosophila melanogaster.
XX
XX OS
XX
XX MO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX N-PSDB; ABL06571.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX
XX Disclosure; SEQ ID NO 14196; 21pp + Sequence listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL20511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
XX ABL22072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_sequences
SQ Sequence 334 AA;
Query Match 13.1%; Score 216.5; DB 4; Length 334;
Best Local Similarity 26.0%; Pred. No. 4.8e-13;
Matches 86; Conservative 48; Mismatches 124; Indels 73; Gaps 9;
QY 3 MYQKPRHGSGAPRLVRLPTCMYRLPNV-----HSKTSPTATDAG----- 42
DB 1 MYELKTL--LPQDIKLPCTCMYRLKNVSLAASGTSASTSASTSSCKLEANNRID 57
QY 43 -----HVOETSBSPLQALRESRODILKRLVELKAAVDGSKMHTPD 84
DB 58 RTGNMATCALDDSLGRQGLLKXDDTASVARQEKVLLKQLELKAQGLQIR----- 110
QY 85 ADDLVNIIQADEPTTLATNTLINSVLGKDYGLAKDIOVINANPASPISLVLVHRLCB 144
DB 111 AGLGVCG--KTFOHTTAFQNG-----GLKEVPLQDVINGHNPFIYPALLALKNAMRN 161
QY 145 RYRVLSVTHHSSVKNV-----PENLYKCRSEQARKSRHEYQGLFTLTKNVKPTQM 197
DB 162 LVTIIVKTFHSTMADIGPAAREFEANLAKVPVNPPLP-----KISVTLIMKNCERTEM 215
QY 198 KFSVQTMCPRIEGBGNIAFLFSLFGQKHNAVTLTL---IDSWDIAIMPQLREGSSKEKA 254
DB 216 ISSPTMYVPIYGEVNIIRVILGRGPAEYRVRBSPCLNEIDVLIDCYQLLRCHTKHTQVA 275
QY 255 VFRSNNASALGRSPWLVGNELTVADVVMISVL 285

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:36:20 ; Search time 17.81 Seconds
(without alignment)
1341.256 Million cell updates/sec

Title: US-10-622-817-5
Perfect score: 1655
Sequence: 1 MPWQVKPHGSGAPLRYEL.....RWKSCENLAPFSTALQLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCUTS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfilee1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------------------|-------------------|
| 1 | 1464 | 88.5 | 341 | 4 US-09-949-016-11312 | Sequence 11312, A |
| 2 | 1436 | 86.8 | 312 | 2 US-08-518-862C-2 | Sequence 2, Appl |
| 3 | 106 | 6.4 | 1512 | 3 US-09-443-184-48 | Sequence 48, Appl |
| 4 | 96.5 | 5.8 | 1440 | 3 US-09-357-251-37 | Sequence 37, Appl |
| 5 | 90.5 | 5.5 | 359 | 3 US-09-540-824-2 | Sequence 2, Appl |
| 6 | 90 | 5.4 | 559 | 2 US-08-756-317-10 | Sequence 10, Appl |
| 7 | 90 | 5.4 | 559 | 4 US-09-091-609-4 | Sequence 4, Appl |
| 8 | 90 | 5.4 | 2954 | 4 US-09-150-867-1 | Sequence 1, Appl |
| 9 | 90 | 5.4 | 5215 | 4 US-09-105-537-2 | Sequence 2, Appl |
| 10 | 89 | 5.4 | 1056 | 4 US-09-595-684B-29 | Sequence 29, Appl |
| 11 | 89 | 5.4 | 1057 | 3 US-09-541-782-10 | Sequence 10, Appl |
| 12 | 89 | 5.4 | 1057 | 4 US-09-723-820-10 | Sequence 10, Appl |
| 13 | 89 | 5.4 | 1057 | 4 US-10-270-085-10 | Sequence 10, Appl |
| 14 | 89 | 5.4 | 1287 | 4 US-09-949-016-7826 | Sequence 7826, Ap |
| 15 | 88 | 5.3 | 712 | 4 US-09-489-039A-10736 | Sequence 10736, A |
| 16 | 88 | 5.3 | 778 | 4 US-09-583-110-3993 | Sequence 3993, Ap |
| 17 | 88 | 5.3 | 786 | 4 US-09-107-433-3893 | Sequence 3893, Ap |
| 18 | 86.5 | 5.2 | 520 | 4 US-09-248-796A-17644 | Sequence 17644, A |
| 19 | 86.5 | 5.2 | 567 | 4 US-09-540-236-2449 | Sequence 2449, Ap |
| 20 | 86 | 5.2 | 935 | 4 US-09-489-039A-11461 | Sequence 11461, A |
| 21 | 85.5 | 5.2 | 556 | 4 US-09-134-000C-6329 | Sequence 6329, Ap |
| 22 | 85.5 | 5.2 | 608 | 4 US-09-284-768A-4 | Sequence 4, Appl |
| 23 | 85.5 | 5.2 | 657 | 4 US-09-284-768A-7 | Sequence 7, Appl |
| 24 | 85 | 5.1 | 222 | 4 US-09-248-796A-18146 | Sequence 18146, A |
| 25 | 85 | 5.1 | 302 | 4 US-09-107-532A-6924 | Sequence 6924, Ap |
| 26 | 85 | 5.1 | 559 | 4 US-09-672-749-2 | Sequence 2, Appl |
| 27 | 85 | 5.1 | 1971 | 4 US-09-914-272A-1 | Sequence 1, Appl |

| | | | | | |
|----|------|-----|------|------------------------|-------------------|
| 28 | 85 | 5.1 | 1971 | 4 US-10-638-333-1 | Sequence 1, Appl |
| 29 | 84 | 5.1 | 443 | 4 US-09-949-016-10598 | Sequence 10598, A |
| 30 | 84 | 5.1 | 646 | 4 US-09-248-796A-18678 | Sequence 18678, A |
| 31 | 84 | 5.1 | 1057 | 4 US-09-428-156B-2 | Sequence 2, Appl |
| 32 | 83.5 | 5.0 | 360 | 4 US-09-949-016-10589 | Sequence 10589, A |
| 33 | 83.5 | 5.0 | 475 | 4 US-09-328-352-5379 | Sequence 5379, Ap |
| 34 | 83.5 | 5.0 | 515 | 4 US-09-489-039A-8208 | Sequence 8208, Ap |
| 35 | 83.5 | 5.0 | 519 | 4 US-09-543-681A-6505 | Sequence 6505, Ap |
| 36 | 83.5 | 5.0 | 789 | 3 US-09-002-285-84 | Sequence 84, Appl |
| 37 | 83.5 | 5.0 | 789 | 4 US-09-589-477-84 | Sequence 84, Appl |
| 38 | 83.5 | 5.0 | 789 | 4 US-10-099-285A-84 | Sequence 84, Appl |
| 39 | 83.5 | 5.0 | 959 | 4 US-09-543-681A-6879 | Sequence 6879, Ap |
| 40 | 83 | 5.0 | 434 | 3 US-09-012-072-4 | Sequence 4, Appl |
| 41 | 83 | 5.0 | 434 | 3 US-09-120-601-4 | Sequence 4, Appl |
| 42 | 83 | 5.0 | 448 | 3 US-09-120-601-6 | Sequence 6, Appl |
| 43 | 83 | 5.0 | 559 | 4 US-09-821-016-1 | Sequence 1, Appl |
| 44 | 83 | 5.0 | 559 | 4 US-10-266-787-1 | Sequence 1, Appl |
| 45 | 83 | 5.0 | 767 | 3 US-08-836-567-8 | Sequence 8, Appl |

ALIGNMENTS

```
RESULT 1
US-09-949-016-11312
; Sequence 11312, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 11312
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11312

Query Match      88.5%; Score 1464; DB 4; Length 341;
Best Local Similarity 87.2%; Pred. No. 96-153;
Matches 279; Conservative 16; Mismatches 25; Indels 0; Gaps 0;

QY      1 MPWQVKPHGSGAPLRYELPTCWRRLPNVHSKTTSPATDAGHVQFTSPSLQLESROD 60
        |||||
DB      22 MPWQVKPHGSGAPLRYELPTCWRRLPNVHSKTTSPATDAGHVQFTSPSLQLESROD 81
        |||||

QY      61 DILKLYELKAADVGSKVIHTPPADIDVTNTIADDEPTLTNTNLSVSKDQALX 120
        |||||
DB      82 DILKLYELKAADVGSKVIHTPPADIDVTNTIADDEPTLTNTNLSVSKDQALX 141
        |||||

QY      121 DIVIANPASPPLSLVLRHLRCERYLVSTYHTSSVKNVDENLVKCGEQAKSRHS 180
        |||||
DB      142 DIVIANPASPPLSLVLRHLRCERYLVSTYHTSSVKNVDENLVKCGEQAKSRHS 201
        |||||

QY      181 YOLGFTLIWKVNPKTQMKFSVOTMCPREGSGNIARFLFSLFCQKNAVTLTLIDSVNDA 240
        |||||
DB      202 YOLGFTLIWKVNPKTQMKFSVOTMCPREGSGNIARFLFSLFCQKNAVTLTLIDSVNDA 261
        |||||

QY      241 MQLREGSSKEKAAVFRSNNSALGRSPMLVGHETLVADVNVLSVYLOOTGSSGAATPNVQ 300
        |||||
DB      262 IQLREGSSKEKAAVFRSNNSALGRSPMLVGHETLVADVNVLSVYLOOTGSSGAATPNVQ 321
        |||||

QY      301 RWKSCENLAPFSTALQLK 320
```

Db 322 RMWRSCEMLAPFNTLKLK 341

RESULT 2

US-08-518-862C-2
 ; Sequence 2, Application US/08518862C
 ; Patent No. 5843757
 ; GENERAL INFORMATION:
 ; APPLICANT: Vogelstein, Bert
 ; APPLICANT: Kinzler, Kenneth W.
 ; APPLICANT: Nicolaides, Nicholas C.
 ; TITLE OF INVENTION: Human JTV1 Gene Overlaps PMS2 Gene
 ; NUMBER OF SEQUENCES: 23
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Banner & Witcoff, Ltd.
 ; STREET: 1001 G Street, N.W.
 ; CITY: Washington, D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20001
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/518,862C
 ; FILING DATE: 24-AUG-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kagan, Sarah A.
 ; REGISTRATION NUMBER: 32,141
 ; REFERENCE/DOCKET NUMBER: 01107.49697
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-508-9100
 ; TELEFAX: 202-508-9299
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 312 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: 1linear
 ; MOLECULE TYPE: protein
 ; US-08-518-862C-2

Query Match 86.8%; Score 1436; DB 2; Length 312;
 Best Local Similarity 87.5%; Pred. No. 9,6e-150;
 Matches 273; Conservative 14; Mismatches 25; Indels 0; Gaps 0;

QY 1 MEMYGVKPYHGGASAPLRYELPTCMYRLPVNHSKTTSPATDAGHVOETSEPSLQALIESROD 60
 DB 1 MEMYGVKPYHGGASAPLRYELPTCMYRLPVNHSKTTSPATDAGHVOETSEPSLQALIESROD 60
 QY 61 DILKRLYEIKAVDGLSKMHTPDADLDVTNLIQADEPTTLATNTLDINSVLTGKDYGLK 120
 DB 61 DILKRLYEIKAVDGLSKMHTPDADLDVTNLIQADEPTTLATNTLDINSVLTGKDYGLK 120
 QY 121 DIVINANASPLSLVYHRLICERYLVSTVHTSSVKNVPEMLVKCFGEQARKQSHE 180
 DB 121 DIVINANASPLSLVYHRLICERYLVSTVHTSSVKNVPEMLVKCFGEQARKQSHE 180
 QY 181 YOLGFTLIWKVPTKQMFVQTMCPRIEGENIARFLPSLFGOGHNAVTLTLIDSWDIA 240
 DB 181 YOLGFTLIWKVPTKQMFVQTMCPRIEGENIARFLPSLFGOGHNAVTLTLIDSWDIA 240
 QY 241 MFOUREGSSKEKAIVFRSMNSALGRSPWLVGNELTVAADVLTMSVLTQOTGSSGAAPTNO 300
 DB 241 MFOUREGSSKEKAIVFRSMNSALGRSPWLVGNELTVAADVLTMSVLTQOTGSSGAAPTNO 300
 QY 301 RWLKSCEMLAPF 312
 DB 301 RWLKSCEMLAPF 312

RESULT 3

US-09-443-184-48
 ; Sequence 48, Application US/09443184A
 ; Patent No. 6372431
 ; GENERAL INFORMATION:
 ; APPLICANT: Cunningham, Mary Jane
 ; APPLICANT: Zweigler, Gary
 ; APPLICANT: Kaser, Matthew R.
 ; APPLICANT: Panzer, Scott
 ; APPLICANT: Selhammer, Jeffrey J.
 ; APPLICANT: Yue, Henry
 ; APPLICANT: Baughn, Mariah
 ; APPLICANT: Azimzai, Yalda
 ; TITLE OF INVENTION: MAMMALIAN TOXICOLOGICAL RESPONSE MARKERS
 ; FILE REFERENCE: PC-0007 US
 ; CURRENT APPLICATION NUMBER: US/09/443,184A
 ; CURRENT FILING DATE: 1999-11-19
 ; NUMBER OF SEQ ID NOS: 138
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 48
 ; LENGTH: 1512
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc,feature
 ; OTHER INFORMATION: Incyte ID No. 6372431 2302721CD1
 ; US-09-443-184-48

Query Match 6.4%; Score 106; DB 3; Length 1512;
 Best Local Similarity 21.8%; Pred. No. 0.096;
 Matches 44; Conservative 33; Mismatches 69; Indels 56; Gaps 8;

QY 124 INANPASPLSLVYHRLICERYLVSTVHTSSVKNVPEMLVKCFGEQARKQSHEYQL 183
 DB 6 LTVNSGDPPLGALL-----AVEHVMD-VSISVEBEKENIILH----- 41
 QY 184 GFTLIWKVPTKQMFVQTMCPRIEGENIARFL-----FSLFGOGHNAVTLTLIDS 237
 DB 42 -----VSENVIFETD-----NSILRYLARVATTAGLYGS--NIMETETIDHWT 82
 QY 238 DIAMFOUREGSSKEKAIVFRSMNSALGRSPWLVGNELTVAADVLTMSVLTQOTGSS 292
 DB 83 EFSATKL--SSCDSFSTINELNHCHLSIRTYLVGNSLSLADLCVWATLKGNAAMQEQLOK 140
 QY 293 GAAPTNOVQWLKSCENLAPST 314
 DB 141 KKAPVHVKRWFGFLERAOQAFQS 162

RESULT 4

US-09-357-251-37
 ; Sequence 37, Application US/09357251
 ; Patent No. 6271441
 ; GENERAL INFORMATION:
 ; APPLICANT: Falco, S. Carl
 ; APPLICANT: Famodu, Layo O.
 ; APPLICANT: Orozco, Buddy
 ; APPLICANT: Schwaber, James S.
 ; TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetase
 ; FILE REFERENCE: BB-1193
 ; CURRENT APPLICATION NUMBER: US/09/357,251
 ; EARLIER FILING DATE: 1999-07-20
 ; EARLIER APPLICATION NUMBER: 60/093,530
 ; NUMBER OF SEQ ID NOS: 37
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 37
 ; LENGTH: 1440
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-357-251-37

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:57:26 ; Search time 50.2443 Seconds
(without alignments)
2084.158 Million cell updates/sec

Title: US-10-622-817-5

Perfect score: 1655

Sequence: 1 MPMYGVKPYHGGSAPLRYEL.....RWLKSCENLAPFSTALQLLK 320

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Published Applications_AA.*
2: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 1436 | 86.8 | 312 | 16 | US-10-463-676-4 |
| 2 | 1069 | 64.6 | 229 | 16 | US-10-463-676-6 |
| 3 | 735 | 44.4 | 161 | 16 | US-10-463-676-5 |
| 4 | 105.5 | 6.4 | 925 | 16 | US-10-437-963-16112 |
| 5 | 104 | 6.3 | 980 | 15 | US-10-369-493-6336 |
| 6 | 103.5 | 6.3 | 468 | 15 | US-10-424-599-26038 |
| 7 | 99.5 | 6.0 | 914 | 16 | US-10-437-963-153870 |
| 8 | 99 | 5.9 | 691 | 15 | US-10-282-122A-44369 |
| 9 | 98 | 5.9 | 690 | 9 | US-09-815-242-5841 |
| 10 | 96.5 | 5.8 | 1440 | 10 | US-09-864-464-37 |
| 11 | 96.5 | 5.8 | 1440 | 16 | US-10-786-720-31 |
| 12 | 96 | 5.8 | 569 | 16 | US-10-437-963-123904 |
| 13 | 95 | 5.7 | 722 | 15 | US-10-369-493-10013 |

| | | | | | | |
|----|------|-----|------|----|----------------------|--------------------|
| 14 | 93.5 | 5.6 | 456 | 15 | US-10-425-114-63166 | Sequence 63166, A |
| 15 | 93 | 5.6 | 204 | 15 | US-10-369-493-10805 | Sequence 10805, A |
| 16 | 93 | 5.6 | 1398 | 16 | US-10-437-963-180124 | Sequence 180124, A |
| 17 | 91 | 5.5 | 201 | 15 | US-10-369-493-13717 | Sequence 13717, A |
| 18 | 91 | 5.5 | 455 | 15 | US-10-425-114-42633 | Sequence 42633, A |
| 19 | 91 | 5.5 | 805 | 15 | US-10-424-599-191796 | Sequence 191796, A |
| 20 | 91 | 5.5 | 1219 | 16 | US-10-437-963-197783 | Sequence 197783, A |
| 21 | 91 | 5.5 | 1788 | 16 | US-10-437-963-197783 | Sequence 197783, A |
| 22 | 90.5 | 5.5 | 419 | 15 | US-10-424-599-260387 | Sequence 260387, A |
| 23 | 90.5 | 5.5 | 578 | 15 | US-10-418-861B-32 | Sequence 32, Appl |
| 24 | 90.5 | 5.5 | 661 | 15 | US-10-369-493-2558 | Sequence 2558, Ap |
| 25 | 90.5 | 5.5 | 716 | 15 | US-10-369-493-2175 | Sequence 2175, Ap |
| 26 | 90 | 5.4 | 559 | 9 | US-09-364-847-21 | Sequence 21, Appl |
| 27 | 90 | 5.4 | 856 | 9 | US-09-364-847-33 | Sequence 33, Appl |
| 28 | 90 | 5.4 | 856 | 9 | US-09-364-847-35 | Sequence 35, Appl |
| 29 | 90 | 5.4 | 1500 | 16 | US-10-437-963-111265 | Sequence 111265, A |
| 30 | 90 | 5.4 | 5215 | 9 | US-09-861-289-2 | Sequence 2, Appl1 |
| 31 | 90 | 5.4 | 5215 | 9 | US-09-860-846-2 | Sequence 2, Appl1 |
| 32 | 90 | 5.4 | 5215 | 10 | US-09-988-364B-2 | Sequence 2, Appl1 |
| 33 | 90 | 5.4 | 5215 | 10 | US-09-836-821-2 | Sequence 2, Appl1 |
| 34 | 90 | 5.4 | 5215 | 14 | US-10-271-889-45 | Sequence 45, Appl |
| 35 | 89.5 | 5.4 | 203 | 15 | US-10-369-493-8604 | Sequence 8604, Ap |
| 36 | 89 | 5.4 | 407 | 15 | US-10-424-599-267113 | Sequence 267113, A |
| 37 | 89 | 5.4 | 1056 | 15 | US-10-282-174-472 | Sequence 472, App |
| 38 | 89 | 5.4 | 1056 | 15 | US-10-282-174-474 | Sequence 474, App |
| 39 | 89 | 5.4 | 1056 | 15 | US-10-282-174-476 | Sequence 476, App |
| 40 | 89 | 5.4 | 1056 | 17 | US-10-600-009-472 | Sequence 472, App |
| 41 | 89 | 5.4 | 1056 | 17 | US-10-600-009-476 | Sequence 476, App |
| 42 | 89 | 5.4 | 1056 | 17 | US-10-600-009-476 | Sequence 476, App |
| 43 | 89 | 5.4 | 1057 | 16 | US-10-714-796-4 | Sequence 4, Appl1 |
| 44 | 88.5 | 5.3 | 656 | 15 | US-10-369-493-18559 | Sequence 18559, A |
| 45 | 88.5 | 5.3 | 799 | 16 | US-10-437-963-112603 | Sequence 112603, A |

ALIGNMENTS

RESULT 1
US-10-463-676-4
Sequence 4, Application US/10463676
GENERAL INFORMATION:
APPLICANT: Kim, Sunghoon
TITLE OF INVENTION: Method for Treating Cancer Using P38/JTV-1 and Method
FILE REFERENCE: 012679-091
CURRENT APPLICATION NUMBER: US/10/463,676
CURRENT FILING DATE: 2003-06-18
PRIOR APPLICATION NUMBER: KR 10-2003-13058
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 312
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (1)...(312)
OTHER INFORMATION: 1-312 amino acid sequence of p38/JTV-1
US-10-463-676-4

Query Match 86.8%; Score 1436; DB 16; Length 312;
Best Local Similarity 87.5%; Pred. No. 2.6e-137;
Matches 273; Conservative 14; Mismatches 25; Indels 0; Gaps 0;
QY 1 MPMYGVKPYHGGSAPLRYELPTCMRLPVVHSGKTSPTADAGVQETSPPSLQALSSROD 60
DB 1 MPMYGVKPYHGGSAPLRYELPTCMRLPVVHSGKTSPTADAGVQETSPPSLQALSSROD 60
QY 61 DILKRLYEIKAVDGLSKMHTPDADLDVNTILQADEPTLATNTLDLNSVLGKDYGALK 120

Db 61 DILKRLYLELKAADVGLSKMIQTDPADLDVTNIIQADEPTTLTNALDLSVYCKDYGALK 120
Qy 121 DIVINANPASPPLSLVLAHRLCEHRRVYSTVHTSSVKNVPENLVKCGEARKQSRHE 180
Db 121 DIVINANPASPPLSLVLAHRLCEHRRVYSTVHTSSVKNVPENLVKCGEARKQSRHE 180
Qy 181 YQGLFTLWKNVKTQMKFSVQTMCPIDEGGNARLFSLFOCKNAATVLTIDSWVIA 240
Db 181 YQGLFTLWKNVKTQMKFSVQTMCPIDEGGNARLFSLFOCKNAATVLTIDSWVIA 240
Qy 241 MPOLREGSSKEKAARVRSNNSALGRSPMLVGNELTVADVILMSVLTQOTGSSGAPITVQ 300
Db 241 IFOLKGGSSKEKAARVRSNNSALGRSPMLVGNELTVADVILMSVLTQOTGSSGAPITVQ 300
Qy 301 RMLKSCENLAPF 312
Db 301 RWRBSCENLAPF 312

RESULT 2
US-10-463-676-6
; Sequence 6; Application US/10463676
; Publication No. US20040175375A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Sunghoon
; APPLICANT: Park, Bum-Joon
; TITLE OF INVENTION: Method for Treating Cancer Using P38/JTV-1 and Method
; FILE REFERENCE: 012679-091
; CURRENT APPLICATION NUMBER: US/10/463,676
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: KR 10-2003-13058
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)...(229)
; OTHER INFORMATION: 84-312 amino acid sequence p38/JTV-1
US-10-463-676-6

Query Match 64.6%; Score 1069; DB 16; Length 229;
Best Local Similarity 88.2%; Pred. No. 4e-100;
Matches 202; Conservative 12; Mismatches 15; Indels 0; Gaps 0;
Qy 84 DADLDVTNIIQADEPTTLTNALDLSVYCKDYGALKDIVINANPASPPLSLVLAHRLC 143
Db 1 DADLDVTNIIQADEPTTLTNALDLSVYCKDYGALKDIVINANPASPPLSLVLAHRLC 60
Qy 144 ERRRVYSTVHTSSVKNVPENLVKCGEARKQSRHEVQGLTLMKNVKTQMKFSVQ 203
Db 61 EHRVYSTVHTSSVKNVPENLVKCGEARKQSRHEVQGLTLMKNVKTQMKFSVQ 120
Qy 204 MCPIDEGGNARLFSLFOCKNAATVLTIDSWVIDAMPOLREGSSKEKAARVRSNNSAL 263
Db 121 MCPIDEGGNARLFSLFOCKNAATVLTIDSWVIDAMPOLREGSSKEKAARVRSNNSAL 180
Qy 264 GRSPMLVGNELTVADVILMSVLTQOTGSSGAPITVQVQ 312
Db 181 GRSPMLVGNELTVADVILMSVLTQOTGSSGAPITVQVQ 229

RESULT 3
US-10-463-676-5
; Sequence 5; Application US/10463676
; Publication No. US20040175375A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Sunghoon
; APPLICANT: Park, Bum-Joon

; TITLE OF INVENTION: Method for Treating Cancer Using P38/JTV-1 and Method
; FILE REFERENCE: 012679-091
; CURRENT APPLICATION NUMBER: US/10/463,676
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: KR 10-2003-13058
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)...(161)
; OTHER INFORMATION: 1-161 amino acid sequence of p38/JTV-1
US-10-463-676-5

Query Match 44.4%; Score 735; DB 16; Length 161;
Best Local Similarity 88.8%; Pred. No. 2.5e-66;
Matches 143; Conservative 5; Mismatches 13; Indels 0; Gaps 0;
Qy 1 MEMVQKHYHGGAPLRYELPTCMYRLPNVHSGKTSPTADAGHVOETSEPSIQALSRQD 60
Db 1 MEMVQKHYHGGAPLRYELPTCMYRLPNVHSGKTSPTADAGHVOETSEPSIQALSRQD 60
Qy 61 DILKRLYLELKAADVGLSKMIHTPDADLDVTNIIQADEPTTLTNALDLSVYCKDYGALK 120
Db 61 DILKRLYLELKAADVGLSKMIHTPDADLDVTNIIQADEPTTLTNALDLSVYCKDYGALK 120
Qy 121 DIVINANPASPPLSLVLAHRLCEHRRVYSTVHTSSVKNV 161
Db 121 DIVINANPASPPLSLVLAHRLCEHRRVYSTVHTSSVKNV 161

RESULT 4
US-10-437-963-126132
; Sequence 126132; Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 126132
; LENGTH: 925
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PRT_MFT4530_28709C.1.pcp
US-10-437-963-126132

Query Match 6.4%; Score 105.5; DB 16; Length 925;
Best Local Similarity 20.3%; Pred. No. 0.51; Indels 99; Gaps 13;
Matches 64; Conservative 46; Mismatches 107;
Qy 25 YLPRVHSGKTSPTADAGHVOETSEPSIQALSRQDI--LKLRLYLELKAADVGLSKMIHT 82
Db 152 YWVDIILKVTGPV-----TIDPRLOALYKKTLELVGIEKSEKLVILSLGDDVHA 203
Qy 83 PDADLDVTNIIQAD--EPTTLATNTLNLNSVYCKDYGA-----LNDIVINAN 127
Db 83 PDADLDVTNIIQAD--EPTTLATNTLNLNSVYCKDYGA-----LNDIVINAN 127

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: February 23, 2005, 13:34:50 ; Search time 13.3213 Seconds
(without alignments)
2311.294 Million cell updates/sec

Title: US-10-622-817-5

Perfect score: 1655

Sequence: 1 MPMYGVKPYHGGSAFLRVEL.....RWLKSCENLAPFTALQLLK 320

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|----------------------|
| 1 | 126.5 | 7.6 | 719 | 2 T52043 | probable glutamate |
| 2 | 112 | 6.8 | 728 | 2 T01200 | probable glutamate |
| 3 | 104 | 6.3 | 980 | 2 T30089 | probable zinc prot |
| 4 | 99 | 6.0 | 691 | 2 B89797 | glycerol ester hyd |
| 5 | 98.5 | 6.0 | 222 | 2 B82441 | probable glutathio |
| 6 | 98.5 | 6.0 | 377 | 2 P69172 | hypothetical prote |
| 7 | 98 | 5.9 | 690 | 2 A24545 | triacylglycerol 11 |
| 8 | 96.5 | 5.8 | 1440 | 1 SYNQT | multifunctional am |
| 9 | 95.5 | 5.8 | 753 | 2 T24869 | hypothetical prote |
| 10 | 95 | 5.7 | 488 | 1 H64113 | cortinoid/iron-sul |
| 11 | 95 | 5.7 | 552 | 2 D82421 | conserved hypothet |
| 12 | 95 | 5.7 | 2655 | 2 D96595 | probable acetyl-Co |
| 13 | 95 | 5.7 | 2471 | 2 T42877 | large tegument pro |
| 14 | 94 | 5.7 | 435 | 2 T01567 | hypothetical prote |
| 15 | 94 | 5.7 | 436 | 2 C86719 | GTP-binding protei |
| 16 | 93.5 | 5.6 | 1064 | 2 B86465 | probable protein k |
| 17 | 93 | 5.6 | 490 | 2 D84999 | ketol-acid reducto |
| 18 | 92 | 5.6 | 754 | 2 S37403 | transcription fact |
| 19 | 91.5 | 5.5 | 436 | 2 T51337 | translational elonga |
| 20 | 91 | 5.5 | 1265 | 2 T47626 | structural mainten |
| 21 | 90.5 | 5.5 | 359 | 2 T37921 | ras-associated pro |
| 22 | 90.5 | 5.5 | 456 | 2 C84764 | hypothetical prote |
| 23 | 90.5 | 5.5 | 642 | 2 T39490 | transcription intl |
| 24 | 90.5 | 5.5 | 716 | 2 T37630 | probable glutamate |
| 25 | 90 | 5.4 | 437 | 2 T51338 | translation elonga |
| 26 | 90 | 5.4 | 559 | 2 A38604 | poly(3-hydroxyalka |
| 27 | 90 | 5.4 | 2954 | 2 T14156 | kinesin-related pr |
| 28 | 89.5 | 5.4 | 317 | 2 T05528 | hypothetical prote |
| 29 | 89.5 | 5.4 | 1597 | 2 S65053 | genome polypeptide |

| | | | | | |
|----|------|-----|------|----------|---------------------|
| 30 | 89.5 | 5.4 | 1714 | 1 S18644 | multifunctional am |
| 31 | 89.5 | 5.4 | 4976 | 2 T14165 | peptide synthetase |
| 32 | 89 | 5.4 | 184 | 2 S74356 | glutathione S-tran |
| 33 | 89 | 5.4 | 763 | 2 AE2443 | penicillin-binding |
| 34 | 89 | 5.4 | 1023 | 2 T31669 | neutral zinc finger |
| 35 | 89 | 5.4 | 1056 | 1 G02157 | kinesin-like spind |
| 36 | 89 | 5.4 | 1247 | 2 A33812 | interphotoreceptor |
| 37 | 88.5 | 5.3 | 602 | 2 AB3542 | GTP-binding protei |
| 38 | 88.5 | 5.3 | 656 | 2 H84206 | acetyl-CoA synthet |
| 39 | 88 | 5.3 | 821 | 2 T24728 | hypothetical prote |
| 40 | 88 | 5.3 | 2279 | 2 T42531 | acetyl-CoA carboxy |
| 41 | 88 | 5.3 | 2280 | 2 T38906 | acetyl-CoA carboxy |
| 42 | 87.5 | 5.3 | 474 | 2 B70597 | hypothetical prote |
| 43 | 87.5 | 5.3 | 820 | 2 H82302 | ATP-dependent heli |
| 44 | 87 | 5.3 | 570 | 2 T15763 | hypothetical prote |
| 45 | 86 | 5.2 | 906 | 2 T45158 | pre-mRNA splicing |

ALIGNMENTS

RESULT 1

T52043 probable glutamate-tRNA ligase (BC 6.1.1.17) [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #ext_change 09-Jul-2004

C:Accession: T52043

R:Day, I.S.; Golovkin, M.; Reddy, A.S.

Biochim. Biophys. Acta 1399, 219-24, 1998

A:Title: Cloning of the cDNA for glutamyl-tRNA synthetase from Arabidopsis thaliana.

A:Reference number: Z24836; MUID:9765600; PMID:9765600

A:Accession: T52043

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-719 <DAY>

A:Cross-references: UNIPROT:082462; EMBL:AF067773; PTD:AA036469.1

A:Superfamily: Yeast glutamate-tRNA ligase; glutamine-tRNA ligase homology

C:Keywords: ligase

Query Match 7.6%; Score 126.5; DB 2; Length 719;

Best Local Similarity 28.0%; Pred. No. 0.017;

Matches 49; Conservative 22; Mismatches 57; Indels 47; Gaps 8;

QY 128 PASPPLSLVLRILLCERVRVLTSTHSSVKNVPENLVKCFGEQAKRSRHEVQGFLL 187

DB 10 PESPPLSLVLRILLCERVRVLTSTHSSVKNVPENLVKCFGEQAKRSRHEVQGFLL 55

QY 188 IKNVPTKQMKFSVQTMCPTEGEGNIARFLFSIFGQKHNAVTLTLIDSWVDIAMPOLREG 247

DB 56 LARVY-----GSAKTLPPFYG--NNAFPSSQIDENVVDYASV-FSSG 94

QY 248 SKEKAANVRNNSALGSPMLVGNELTVADYVLSVLOQTGGSGAAPTNQWRW 302

DB 95 SEFENAC--GRVDKYLSSSTFLVGHSLSDYDAVWALAGTG-----QRW 137

RESULT 2

T01200 probable glutamate-tRNA ligase (BC 6.1.1.17) F21E10.12 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #ext_change 09-Jul-2004

C:Accession: T01200

R:Davidson, S.; Rohlfing, T.; David, M.; O'Brian, D.

submitted to the EMBL Data Library, April 1998

A:Description: The sequence of A. thaliana F21E10.

A:Reference number: Z14258

A:Accession: T01200

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-728 <DAY>

A:Cross-references: UNIPROT:065253; EMBL:AF058914; NID:93047074; PTD:93047084; GSPDB:GNC

A:Experimental source: cultivar Columbia

C:Genetics:

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:33:04 ; Search time 60.8145 Seconds
(without alignments)
2694.512 Million cell updates/sec

Title: US-10-622-817-5

Perfect score: 1655

Sequence: 1 MPMYGVKPYHGGAPLRLV.....RLKSCENLAPSTALQLLK 320

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|--------------|--------------------|
| 1 | 1655 | 100.0 | 320 | Q8R010 | Q8R010 mus musculu |
| 2 | 1652 | 99.8 | 320 | Q8RZY6 | Q8RZY6 mus musculu |
| 3 | 1573 | 95.0 | 320 | 1 MCA2_CRIGR | Q9WY77 cricetus |
| 4 | 1464 | 88.5 | 320 | 1 MCA2_HUMAN | Q13155 homo sapien |
| 5 | 1460 | 88.2 | 320 | 2 Q96CZ5 | Q96CZ5 mus musculu |
| 6 | 1412 | 85.3 | 280 | 2 Q8R3V2 | Q8R3V2 mus musculu |
| 7 | 1153.5 | 69.7 | 311 | 2 Q6DK86 | Q6DK86 xenopus tro |
| 8 | 1146.5 | 69.3 | 311 | 2 Q6IND4 | Q6IND4 xenopus lae |
| 9 | 1146.5 | 69.3 | 311 | 2 Q7ZYD7 | Q7ZYD7 xenopus lae |
| 10 | 926.5 | 56.0 | 321 | 2 Q7J3C0 | Q7J3C0 brachydanto |
| 11 | 247 | 14.9 | 340 | 2 Q7Q7A3 | Q7Q7A3 anopheles g |
| 12 | 224 | 13.5 | 301 | 2 Q7KUM5 | Q7KUM5 drosophila |
| 13 | 224 | 13.5 | 313 | 2 Q8T060 | Q8T060 drosophila |
| 14 | 224 | 13.5 | 322 | 2 Q6NKM4 | Q6NKM4 drosophila |
| 15 | 216.5 | 13.1 | 334 | 1 MCA2_DROME | Q9VUR3 drosophila |
| 16 | 126.5 | 7.6 | 719 | 2 Q82462 | Q82462 arabidopsis |
| 17 | 112 | 6.8 | 728 | 2 Q65253 | Q65253 arabidopsis |
| 18 | 107 | 6.5 | 715 | 2 Q9LIZ8 | Q9LIZ8 oryza sativ |
| 19 | 106 | 6.4 | 328 | 2 Q86X73 | Q86X73 homo sapien |
| 20 | 106 | 6.4 | 559 | 2 Q8RPZ6 | Q8RPZ6 pseudomona |
| 21 | 106 | 6.4 | 869 | 2 Q6P577 | Q6P577 homo sapien |
| 22 | 106 | 6.4 | 1721 | 2 Q7NID9 | Q7NID9 glisobacter |
| 23 | 104 | 6.3 | 980 | 2 Q17592 | Q17592 caenorhabd |
| 24 | 102.5 | 6.2 | 420 | 2 Q8PQ00 | Q8PQ00 coxynabacte |
| 25 | 102 | 6.2 | 559 | 2 Q939A8 | Q939A8 pseudomona |
| 26 | 102 | 6.2 | 1048 | 2 Q7N3E2 | Q7N3E2 photorhabd |
| 27 | 101.5 | 6.1 | 715 | 2 Q91B77 | Q91B77 white spot |
| 28 | 101.5 | 6.1 | 730 | 2 Q91BF5 | Q91BF5 white spot |
| 29 | 101.5 | 6.1 | 913 | 2 Q6DRB3 | Q6DRB3 brachydanto |
| 30 | 101.5 | 6.1 | 913 | 2 Q6PEQ4 | Q6PEQ4 brachydanto |
| 31 | 100.5 | 6.1 | 555 | 2 Q7MEV7 | Q7MEV7 vibrio vuln |

| | | | | | | |
|----|-------|-----|------|---|-----------|--------------------|
| 32 | 100.5 | 6.1 | 555 | 2 | Q8D7U3 | Q8D7U3 vibrio vuln |
| 33 | 100.5 | 6.1 | 922 | 2 | Q8K2J4 | Q8K2J4 mus musculu |
| 34 | 99.5 | 6.0 | 885 | 2 | Q6TEM5 | Q6TEM5 brachydanto |
| 35 | 99 | 6.0 | 559 | 2 | Q8R067 | Q8R067 pseudomona |
| 36 | 99 | 6.0 | 690 | 1 | LIP STRAW | Q8Y622 staphylococ |
| 37 | 99 | 6.0 | 690 | 2 | Q6GCF1 | Q6GCF1 staphylococ |
| 38 | 99 | 6.0 | 691 | 2 | Q79827 | Q79827 staphylococ |
| 39 | 99 | 6.0 | 691 | 2 | Q99WQ6 | Q99WQ6 staphylococ |
| 40 | 99 | 6.0 | 691 | 2 | Q7A7P2 | Q7A7P2 staphylococ |
| 41 | 98.5 | 6.0 | 222 | 2 | Q9KM05 | Q9KM05 vibrio chol |
| 42 | 98.5 | 6.0 | 377 | 2 | Q26161 | Q26161 methanobact |
| 43 | 98.5 | 6.0 | 1597 | 2 | Q8V9A2 | Q8V9A2 crucifer to |
| 44 | 98 | 5.9 | 240 | 2 | Q8BTR1 | Q8BTR1 mus musculu |
| 45 | 98 | 5.9 | 690 | 1 | LIP STRAW | P10335 staphylococ |

ALIGNMENTS

RESULT 1
Q8R010 PRELIMINARY; PRT; 320 AA.
ID Q8R010
AC Q8R010;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Jtv1-pending protein.
GN Name=Jtv1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N, and FVB/N-3; TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ussid T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny K.C., Harte S., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley A.C., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez R.A., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
RT Jones S.J., Maier M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N, TISSUE=Mammary tumor. C3;
RA Strausberg R.;
RN Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N-3; TISSUE=Mammary tumor;
RA Strausberg R.;
RN Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL, BC024410; AAH24410.1; -
DR EMBL, BC026972; AAH26972.1; -
DR MGD, MGI:2385237; Jtv1.
DR InterPro, IPR004046; GST_Cterm.
DR InterPro, IPR010987; GST_C_1ike.
DR Pfam, PF00043; GST_C_1.
SQ SEQUENCE 320 AA; 35396 MW; 1745D7E4BC3670D CRC64;

Query Match 100.0%; Score 1655; DB 2; Length 320;
 Best Local Similarity 100.0%; Pred. No. 2.6e-128;
 Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPWYQKPYHGGASAPLRVELPTCMYRLPNVHSKTTSPATDAGHVGETSEPSLQALSRQD 60
 DB 1 MPWYQKPYHGGASAPLRVELPTCMYRLPNVHSKTTSPATDAGHVGETSEPSLQALSRQD 60
 QY 61 DILKRLYELKAADVGSKNIHTPPADLDVTNIIQADEPTTLATNTLDNSVLGXDYGALK 120
 DB 61 DILKRLYELKAADVGSKNIHTPPADLDVTNIIQADEPTTLATNTLDNSVLGXDYGALK 120
 QY 121 DIVINANPASPLSLVLRLLCERYRLSTVHTSSVKNVPELVKCFGEQARKSRHE 180
 DB 121 DIVINANPASPLSLVLRLLCERYRLSTVHTSSVKNVPELVKCFGEQARKSRHE 180
 QY 121 DIVINANPASPLSLVLRLLCERYRLSTVHTSSVKNVPELVKCFGEQARKSRHE 180
 DB 121 DIVINANPASPLSLVLRLLCERYRLSTVHTSSVKNVPELVKCFGEQARKSRHE 180
 QY 181 YOLGFTLLWKVNPKTQMKFSVQTMCPLEGEGNIARFLPSLFGQKHNAVTLTLIDSWDIA 240
 DB 181 YOLGFTLLWKVNPKTQMKFSVQTMCPLEGEGNIARFLPSLFGQKHNAVTLTLIDSWDIA 240
 QY 241 MFOLREGSSKEKAIVFSSMSALGRSPMLVGNELTVADVLMVSLQOTGSSGGAAPTNVQ 300
 DB 241 MFOLREGSSKEKAIVFSSMSALGRSPMLVGNELTVADVLMVSLQOTGSSGGAAPTNVQ 300
 QY 301 RWLKSCEMLAPFSTALQLLK 320
 DB 301 RWLKSCEMLAPFSTALQLLK 320

RESULT 2
 QY 0QR2Y6 PRELIMINARY; PRT; 320 AA.
 AC 0QR2Y6;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Jty1-Pending protein.
 GN Name=Jty1;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCB1_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CZECH II; TISSUE=Mammary tumor;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Mak S.T., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gumarathne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
 RA Villalon D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallos D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CZECH II; TISSUE=Mammary tumor;
 RX Strausberg R.;
 Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL, BC026958; AAH26958.1; -.

DR MGI:2385237; Jty1.
 DR InterPro: IPR004046; GST_Cterm.
 DR InterPro: IPR010987; GST_C1like.
 DR Pfam: PF00043; GST_C; 1.
 SQ SEQUENCE 320 AA; 35423 MW; 1C21FLA74C9882B4 CRC64;

Query Match 99.8%; Score 1652; DB 2; Length 320;
 Best Local Similarity 99.7%; Pred. No. 4.6e-128;
 Matches 319; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPWYQKPYHGGASAPLRVELPTCMYRLPNVHSKTTSPATDAGHVGETSEPSLQALSRQD 60
 DB 1 MPWYQKPYHGGASAPLRVELPTCMYRLPNVHSKTTSPATDAGHVGETSEPSLQALSRQD 60
 QY 61 DILKRLYELKAADVGSKNIHTPPADLDVTNIIQADEPTTLATNTLDNSVLGXDYGALK 120
 DB 61 DILKRLYELKAADVGSKNIHTPPADLDVTNIIQADEPTTLATNTLDNSVLGXDYGALK 120
 QY 121 DIVINANPASPLSLVLRLLCERYRLSTVHTSSVKNVPELVKCFGEQARKSRHE 180
 DB 121 DIVINANPASPLSLVLRLLCERYRLSTVHTSSVKNVPELVKCFGEQARKSRHE 180
 QY 121 DIVINANPASPLSLVLRLLCERYRLSTVHTSSVKNVPELVKCFGEQARKSRHE 180
 DB 121 DIVINANPASPLSLVLRLLCERYRLSTVHTSSVKNVPELVKCFGEQARKSRHE 180
 QY 181 YOLGFTLLWKVNPKTQMKFSVQTMCPLEGEGNIARFLPSLFGQKHNAVTLTLIDSWDIA 240
 DB 181 YOLGFTLLWKVNPKTQMKFSVQTMCPLEGEGNIARFLPSLFGQKHNAVTLTLIDSWDIA 240
 QY 241 MFOLREGSSKEKAIVFSSMSALGRSPMLVGNELTVADVLMVSLQOTGSSGGAAPTNVQ 300
 DB 241 MFOLREGSSKEKAIVFSSMSALGRSPMLVGNELTVADVLMVSLQOTGSSGGAAPTNVQ 300
 QY 301 RWLKSCEMLAPFSTALQLLK 320
 DB 301 RWLKSCEMLAPFSTALQLLK 320

RESULT 3
 MCA2_CRIGR STANDARD; PRT; 320 AA.
 ID MCA2_CRIGR
 AC 09WVW7;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Multisynthetase complex auxiliary component p38.
 OS Cricetus griseus (Chinese hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Cricetus.
 NC NCB1_TaxID=10029;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RX MEDLINE=9906915; PubMed=9878398; DOI=10.1006/jmbi.1998.2316;
 RA Quevillon S., Robinson J.-C., Berthomieu E., Slatkova M., Mirande M.,
 RT "Molecular assembly of aminoacyl-tRNA synthetases:
 identification of protein-protein interactions and characterization of
 a core protein.";
 RL J. Mol. Biol. 285:183-195(1999).
 RT -1- FUNCTION: Probable core protein of the multisynthetase complex
 that serves as a template for the assembly of the supramolecular
 structure.
 CC -1- SUBUNIT: Component of the multisynthetase complex which is
 composed of a bifunctional glutamyl-prolyl-tRNA synthetase, the
 monospecific isoleucyl, leucyl, glutamyl, methionyl, lysyl,
 arginyl, and aspartyl-tRNA synthetases as well as three auxiliary
 proteins, p18, p48 and p43.
 CC -1- SIMILARITY: Contains 1 GST-like domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.ebi.ac.uk/announcement/>)

or send an email to license@sb-sib.ch.

CC EMBL; AF072727; AAD38422.1; -
 CC DR InterPro; IPR010987; GST_C_1like.
 DR pfam; PF004046; GST_C_1.
 DR Protein biosynthesis.
 KW SEQUENCE 320 AA; 35433 MW; 6D24E033ABEC810A CRC64;

Query Match 95.0%; Score 1573; DB 1; Length 320;
 Best Local Similarity 94.4%; Pred. No. 1.6e-121;
 Matches 302; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 1 MEMYQVYPHGGAPRLVRLPTCMYRLPNVHSTKTSPPADAGHVGQETSEPSIQALERSRD 60
 DB 1 MEMYQVKSHYGGAPRLVRLPTCMYRLPNVHSTKTSPPADAGHVGQETSEPSIQALERSRD 60
 QY 61 DILKRLYELKAAVDGSKMHPDADLDVTNIIQADEPTTLATNTLDLSVGLKDYGALK 120
 DB 61 DILKRLYELKAAVDGSKMHPDADLDVTNIIQADEPTTLATNTLDLSVGLKDYGALK 120
 QY 121 DIVINANPASPPLSLVLRLLCERRVLTSTVHTSSVKNVENVKCTGEQARKQSRHE 180
 DB 121 DIVINANPASPPLSLVLRLLCERRVLTSTVHTSSVKNVENVKCTGEQARKQSRHE 180
 QY 181 YQLGFTLWKQVPTQMKFSVOTMCPIDEGGNARFLFSLFGQKHAIVTLTIDSVVDIA 240
 DB 181 YQLGFTLWKQVPTQMKFSVOTMCPIDEGGNARFLFSLFGQKHAIVTLTIDSVVDIA 240
 QY 241 MEQLREGSGSKKAAVFRSMNSALGRSPMLVGNELTVADVLVLSVLTQOTGGSGGAFTNYQ 300
 DB 241 MEQLREGSGSKKAAVFRSMNSALGRSPMLVGNELTVADVLVLSVLTQOTGGSGGAFTNYQ 300
 QY 301 RMLKSCENLAPSTALQLLK 320
 DB 301 RMLKSCENLAPSTALQLLK 320

RESULT 4

KCA2_HUMAN STANDARD; PRT; 320 AA.
 ID Q13155; Q9PIL2;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Multisynthetase complex auxiliary component p38 (JTV-1 protein)
 DE (PRO0992).
 GN Name=JTV1;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OC NCBI_TaxId=9606;
 RN 1;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96115582; PubMed=8666379;
 RA Nicolson J.D., Kinsler K.W., Vogelstein B.;
 RT "Analysis of the 5' region of PMS2 reveals heterogeneous transcripts
 and a novel overlapping gene.";
 RL Genomics 29:329-334 (1995).
 RT 12;
 RC SEQUENCE FROM N.A.
 RP TISSUE=lymph;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strauberg R.L., Feigold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer T.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.U., Uediri T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Millaly S.J.,
 RA Boeck S.A., Mcwan P.J., McKernan K.J., Melek J.A., Guarnante P.H.,
 RA Richard S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN (3)
 RP SEQUENCE OF 197-320 FROM N.A.
 RC TISSUE=Fetal liver;
 RA Zhang C., Yu Y., Zhang S., Wei H., Zhou G., Ouyang S., Luo L., Bi J.,
 RA Liu M., He F.;
 RT "Functional prediction of the coding sequences of 121 new genes
 RT deduced by analysis of cDNA clones from human fetal liver.";
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBA database.
 RN (4)

RP INTERACTION WITH FUBP1.
 RX MEDLINE=22716800; PubMed=12819782; DOI=10.1038/ng1182;
 RA Kim M.J., Park B.-J., Kang Y.-S., Kim H.J., Park J.-H., Kang J.W.,
 RA Lee S.W., Han J.M., Lee H.-W., Kim S.;
 RT "Downregulation of FUSE-binding protein and c-myc by cRNA synthetase
 RT cofactor p38 is required for lung cell differentiation.";
 RL Nat. Genet. 34:330-336 (2003).

CC -1- FUNCTION: Probable core protein of the multisynthetase complex
 CC that serves as a template for the assembly of the supramolecular
 CC structure. Mediates ubiquitination of FUBP1 and its degradation by
 CC the proteasome.

CC -1- SUBUNIT: Component of the multisynthetase complex which is
 CC comprised of a bifunctional glutamyl-prolyl-cRNA synthetase, the
 CC monospecific isoleucyl, leucyl, glutamyl, methionyl, lysyl,
 CC arginyl, and aspartyl-cRNA synthetases as well as three auxiliary
 CC proteins, p18, p48 and p43. Binds FUBP1.

CC -1- SIMILARITY: Contains 1 GST-like domain.
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
 CC frameshift in position 312.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.sib-sb.ch/announce/>
 CC or send an email to license@sib-sb.ch).

CC EMBL; U24169; AAC50391.1; ALT FRAME.

DR EMBL; BC002853; AAH02853.1; -

DR EMBL; BC010156; AAH010156.1; -

DR EMBL; AF116151; AAF71039.1; -

DR H-InvDB; HIX0006460; -

DR MIM; 600859; -

DR InterPro; IPR010987; GST_C_1like.

DR InterPro; IPR004046; GST_Cterm.

DR pfam; PF00043; GST_C_1.

KW Protein biosynthesis.

KW SEQUENCE 320 AA; 35349 MW; F253726B63C12BAB CRC64;

Query Match 88.5%; Score 1464; DB 1; Length 320;
 Best Local Similarity 87.2%; Pred. No. 1.6e-112;
 Matches 279; Conservative 16; Mismatches 25; Indels 0; Gaps 0;

QY 1 MEMYQVYPHGGAPRLVRLPTCMYRLPNVHSTKTSPPADAGHVGQETSEPSIQALERSRD 60
 DB 1 MEMYQVYPHGGAPRLVRLPTCMYRLPNVHSTKTSPPADAGHVGQETSEPSIQALERSRD 60
 QY 61 DILKRLYELKAAVDGSKMHPDADLDVTNIIQADEPTTLATNTLDLSVGLKDYGALK 120
 DB 61 DILKRLYELKAAVDGSKMHPDADLDVTNIIQADEPTTLATNTLDLSVGLKDYGALK 120
 QY 121 DIVINANPASPPLSLVLRLLCERRVLTSTVHTSSVKNVENVKCTGEQARKQSRHE 180
 DB 121 DIVINANPASPPLSLVLRLLCERRVLTSTVHTSSVKNVENVKCTGEQARKQSRHE 180

```
Dh 121 DIVINANPASPLSLVLRLLCEHVRVLTSTVHTSSVKSVPENLKKCFGEONKKQPPROD 180
Qy 181 YOLGFTLLIKWNPVKTKQMKFSVOTMCPPIBEGNIAFLFSLFGQKNAVTLTLDISWDIA 240
Db 181 YOLGFTLLIKWNPVKTKQMKFSVOTMCPPIBEGNIAFLFSLFGQKNAVTLTLDISWDIA 240
Qy 241 MFOLREGSKERKAAPFRSNNSALGSRPWLVGNELVADVVMVMSVLQQTGGSSGAAPTNVQ 300
Db 241 IFOLKESGSKERKAAPFRSNNSALGSRPWLAGNELTVADVVMVMSVLQQTGGSSVTPVANVQ 300
Qy 301 RWLKSCENLAPFSTALQILK 320
Db 301 RWRMSCENLAPFSTALQILK 320

RESULT 5
096CZ5 PRELIMINARY; PRT; 320 AA.
AC 096CZ5;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE JTV1.
GN Name=JTV1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Uedin T.B., Toshlyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Miliaty S.J.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalka U., Smalhus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC013630; AAH13630.1;
DR InterPro: IPR004046; GST_Cterm.
DR InterPro: IPR010987; GST_C_1like.
DR Pfam: PF00043; GST_C_1.
SQ SEQUENCE 320 AA; 35335 MW; 19F14BF758612E08 CRC64;

Query Match 88.2%; Score 1460; DB 2; Length 320;
Best Local Similarity 86.9%; Pred. No. 3,4e-112;
Matches 278; Conservative 16; Mismatches 26; Indels 0; Gaps 0;

Qy 1 MEMYGVKPYHGGSAALRYELPTCMYRLPNVSHSKTSPATDAGHVOETSEPSLQALLESROD 60
Db 1 MEMYGVKPYHGGSAALRYELPTCMYRLPNVSHSKTSPATDAGHVOETSEPSLQALLESROD 60
Qy 61 DILKRLYEKKAADVGLSKMIQTDPADLDVTNIIQADEPTTLTNALDINSVLGKDYALK 120
Db 61 DILKRLYEKKAADVGLSKMIQTDPADLDVTNIIQADEPTTLTNALDINSVLGKDYALK 120
```

```
Dh 61 DILKRLYEKKAADVGLSKMIQTDPADLDVTNIIQADEPTTLTNALDINSVLGKDYALK 120
Qy 121 DIVINANPASPLSLVLRLLCEHVRVLTSTVHTSSVKSVPENLKKCFGEONKKQPPROD 180
Db 121 DIVINANPASPLSLVLRLLCEHVRVLTSTVHTSSVKSVPENLKKCFGEONKKQPPROD 180
Qy 181 YOLGFTLLIKWNPVKTKQMKFSVOTMCPPIBEGNIAFLFSLFGQKNAVTLTLDISWDIA 240
Db 181 YOLGFTLLIKWNPVKTKQMKFSVOTMCPPIBEGNIAFLFSLFGQKNAVTLTLDISWDIA 240
Qy 241 MFOLREGSKERKAAPFRSNNSALGSRPWLVGNELVADVVMVMSVLQQTGGSSGAAPTNVQ 300
Db 241 IFOLKESGSKERKAAPFRSNNSALGSRPWLAGNELTVADVVMVMSVLQQTGGSSVTPVANVQ 300
Qy 301 RWLKSCENLAPFSTALQILK 320
Db 301 RWRMSCENLAPFSTALQILK 320

RESULT 6
09R3V2 PRELIMINARY; PRT; 280 AA.
AC 09R3V2;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE JTV1.
GN Name=JTV1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Uedin T.B., Toshlyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Miliaty S.J.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalka U., Smalhus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC024480; AAH24480.1;
DR MGI: MGI:2385237; JTV1.
DR InterPro: IPR004046; GST_Cterm.
DR InterPro: IPR010987; GST_C_1like.
DR Pfam: PF00043; GST_C_1.
SQ SEQUENCE 280 AA; 31097 MW; E5BE81498983FD2E CRC64;

Query Match 85.3%; Score 1412; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 2,6e-108;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 QETSEPSLQALLESRODILKRLYEKKAADVGLSKMIQTDPADLDVTNIIQADEPTTLTN 104
Db 45 QETSEPSLQALLESRODILKRLYEKKAADVGLSKMIQTDPADLDVTNIIQADEPTTLTN 104
```



```
DB 5 OETSEPSLOALBESRQODILKRLYELKAAVDGSKMHTPADVDVNNILOADPPTLATN 64
105 TLIDNLVIGKDVGLADIVYNANPASPSTLVLHRLCERYVLSVTHHSSVKNVPPEN 164
DB 65 TLIDNLVIGKDVGLADIVYNANPASPSTLVLHRLCERYVLSVTHHSSVKNVPPEN 124
QY 165 LVKCFEQAOKOSRHEVQGLFTLIWKVNPPTQMKFSVQTMCPTEGEGNIARFLFSGOK 224
DB 125 LVKCFEQAOKOSRHEVQGLFTLIWKVNPPTQMKFSVQTMCPTEGEGNIARFLFSGOK 184
QY 225 HNAVTLTLIDSVWDIAMPOLREGSSKEKAAPFRSNMNSALGRSPMLVGNELTVADVVLMSV 284
DB 185 HNAVTLTLIDSVWDIAMPOLREGSSKEKAAPFRSNMNSALGRSPMLVGNELTVADVVLMSV 244
QY 285 LOOTGSSGAAPPTNVORWLKSCENLAPFSTALQLLK 320
DB 245 LOOTGSSGAAPPTNVORWLKSCENLAPFSTALQLLK 280

RESULT 7
Q6DK86 PRELIMINARY; PRT; 311 AA.
AC Q6DK86;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE MGCG9221 protein.
GN Name=MGCG9221;
OS Xenopus tropicalis (western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
NCBI_Taxid=8364;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Tohyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL, BC074561, AAH74561.1;
DR InterPro, IPR004046, GST_Cterm.
DR InterPro, IPR010987, GST_C_1like.
DR Pfam, PF00043, GST_C_1.
SO SEQUENCE 311 AA, 34480 MW, D98F27F73C466154 CRC64;
```

```
Query Match 69.7%; Score 1153.5; DB 2; Length 311;
Best Local Similarity 67.5%; Pred. No. 6, 8e-87;
Matches 216; Conservative 49; Mismatches 46; Indels 9; Gaps 3;
```

```
QY 1 MPMYQVYPYHGSAPLRLVELPTCMYRLPNVHASKTSPATDAGHVOETSEPSLOALBESROD 60
DB 1 MPMYQVYPYHGSAPLRLVELPTCMYRLPNVHASKTSPATDAGHVOETSEPSLOALBESROD 52
QY 61 DILKRLYELKAAVDGSKMHTPADVDVNNILOADPPTLATN1DLNLSVLCKYGAALK 120
DB 53 DILKRLYELKAAVDGSKMHTPADVDVNNILOADPPTLATN1DLNLSVLCKYGAALK 112
QY 121 DIVYNANPASPSTLVLHRLCERYVLSVTHHSSVKNVPPENLVKCFEQAOKOSRHE 180
DB 113 DIVYNANPASPSTLVLHRLCERYVLSVTHHSSVKNVPPENLVKCFEQAOKOSRHE 172
QY 181 YQGLFTLIWKVNPPTQMKFSVQTMCPTEGEGNIARFLFSGOKHNAVTLTLIDSVWDIA 240
DB 173 YQGLFTLIWKVNPPTQMKFSVQTMCPTEGEGNIARFLFSGOKHNAVTLTLIDSVWDIA 232
QY 241 MFOLREGSSKEKAAPFRSNMNSALGRSPMLVGNELTVADVVLMSVLOOTGSSGAAPPTNVQ 300
DB 233 IFOLREGSSKEKAAPFRSNMNSALGRSPMLVGNELTVADVVLMSVLOOTGSSGAAPPTNVQ 291
QY 301 RMLKSCENLAPFSTALQLLK 320
DB 292 RMLKSCENLAPFSTALQLLK 311

RESULT 8
Q6IND4 PRELIMINARY; PRT; 311 AA.
AC Q6IND4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MGCG80304 protein.
GN Name=MGCG80304;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
NCBI_Taxid=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Tohyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
```

RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Klein S., Gerhard D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL, BC072178, AAH72178.1;
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR010987; GST_C_1like.
DR Pfam; PF00043; GST_C_1.
SQ SEQUENCE 311 AA, 34424 MW, D5E8325C18D88751 CRC64;

Query Match 69.3%; Score 1146.5; DB 2, Length 311;
Best Local Similarity 68.1%; Pred. No. 2.6e-86;
Matches 218; Conservative 44; Mismatches 49; Indels 9; Gaps 3;

QY 1 MPMTQVKEPHGSGAPLRYELPTCMYRLPNVSKTTPATDAGHVOETSEPSLOAESRQD 60
DB 1 MPMTQVKEPHGSGA...IQVDLPTCMYRLPNVSKTTPATDAGHVOETSEPSLOAESRQD 52
QY 61 DILKRIYELKAAVGLSKMTHTPPADLDVNTIIQADEPTLATNTLDINSVLGKYGALK 120
DB 53 DILKRIYELKAAVGLSKMTHTPPADLDVNTIIQADEPTLATNTLDINSVLGKYGALK 112
QY 121 DIVINAPASPLSLVLRHLRCERYVLTSTVTHSSVKNVPELVKCFGEQARKOSRHE 180
DB 113 DIVINAPASPLSLVLRHLRCERYVLTSTVTHSSVKNVPELVKCFGEQARKOSRHE 172
QY 181 YOLGFTLIMKNVPTQKRFVQTCPIEGEGNIGRFLFSIFGQKHNAAVTLTIDSWVDIA 240
DB 173 YOLGFTLIMKNVPTQKRFVQTCPIEGEGNIGRFLFSIFGQKHNAAVTLTIDSWVDIA 232
QY 241 MFOLREGSSKEKAAVFRSNMNSALGRSPMLVGNELTVADVVLMSVLTQOTGSSGAAPTNO 300
DB 233 IFOLREGSSKEKAAVFRSNMNSALGRSPMLVGNELTVADVVLMSVLTQOTGSSGAAPTNO 291
QY 301 RWLKSCENLAPFSTALQLK 320
DB 292 KMKKSCENLASFKSVLRFLK 311

RESULT 9

ID Q7ZXD7 PRELIMINARY; PRT; 311 AA.
AC Q7ZXD7;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Utl-v1-prov protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheef C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Cavatoni T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.J.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Krzywinski M.I., Skalska U., Smallus D.B., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]

RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.,
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT Initiative";
RL Dev. Dyn. 225:384-391 (2002).
RN [3]

RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.,
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL, BC043832, AAH43832.1;
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR010987; GST_C_1like.
DR Pfam; PF00043; GST_C_1.
SQ SEQUENCE 311 AA, 34647 MW, 2F08C9DA60DD63BE CRC64;

Query Match 69.3%; Score 1146.5; DB 2, Length 311;
Best Local Similarity 68.1%; Pred. No. 2.6e-86;
Matches 218; Conservative 44; Mismatches 49; Indels 9; Gaps 4;

QY 1 MPMTQVKEPHGSGAPLRYELPTCMYRLPNVSKTTPATDAGHVOETSEPSLOAESRQD 60
DB 1 MPMTQVKEPHGSGA...IQVDLPTCMYRLPNVSKTTPATDAGHVOETSEPSLOAESRQD 52
QY 61 DILKRIYELKAAVGLSKMTHTPPADLDVNTIIQADEPTLATNTLDINSVLGKYGALK 120
DB 53 DILKRIYELKAAVGLSKMTHTPPADLDVNTIIQADEPTLATNTLDINSVLGKYGALK 112
QY 121 DIVINAPASPLSLVLRHLRCERYVLTSTVTHSSVKNVPELVKCFGEQARKOSRHE 180
DB 113 DIVINAPASPLSLVLRHLRCERYVLTSTVTHSSVKNVPELVKCFGEQARKOSRHE 172
QY 181 YOLGFTLIMKNVPTQKRFVQTCPIEGEGNIGRFLFSIFGQKHNAAVTLTIDSWVDIA 240
DB 173 YOLGFTLIMKNVPTQKRFVQTCPIEGEGNIGRFLFSIFGQKHNAAVTLTIDSWVDIA 232
QY 241 MFOLREGSSKEKAAVFRSNMNSALGRSPMLVGNELTVADVVLMSVLTQOTGSSGAAPTNO 300
DB 233 IFOLREGSSKEKAAVFRSNMNSALGRSPMLVGNELTVADVVLMSVLTQOTGSSGAAPTNO 291
QY 301 RWLKSCENLAPFSTALQLK 320
DB 292 KMKKSCENLASFKSVLRFLK 311

RESULT 10

ID Q7T3C0 PRELIMINARY; PRT; 321 AA.
AC Q7T3C0;
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE zgc:63976.
GN ORFNames=zgc:63976;
OS Brachydanio rerio (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheef C.F., Bhat N.K.,

RA Hopkine R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stajichon M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.B.,
 RA Brownstein M.J., Usdin T.B., Toehlyuk S., Carninci P., Prange C.,
 RA Raha S.B., Loquellano N.A., Peters G.J., Abraham R.D., Mullany S.J.,
 RA Borek S.A., McEwan P.J., McKernan K.U., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettlemen M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smaltz D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN (12)
 RP SEQUENCE FROM N.A.
 RC Tissue=Kidney;
 RA Straube R.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL, BC051178, AAH51178.1; -
 DR ZFIN, ZDB-GENE-040426-2652; zgc:63976.
 DR InterPro, IPR004046; GST_Cterm.
 DR InterPro, IPR010987; GST_C_1like.
 DR Pfam, PF00043; GST_C_1.
 SQ SEQUENCE 321 AA; 34852 MW; BA6F6B951208244A CRC64;

Query Match 56.0%; Score 926.5; DB 2; Length 321;
 Best Local Similarity 57.3%; Pred. No. 4.1e-68;
 Matches 188; Conservative 48; Mismatches 77; Indels 15; Gaps 7;

QY 1 MPMYGVKRYHGSAPLRLVELPTCMYRLPNVHASKTSP---PATDAGHYQETSPPSLOAES 57
 DB 1 MPMYGVKRYHGSAPLRLVELPTCMYRLPNVHASKTSP---PATDAGHYQETSPPSLOAES 54
 QY 58 RODDLIKRLYEKAAVLDGSLKMIHTPPADLDVTNI---LQADEPTLTNTLDLSVLGK 114
 DB 55 RODELLIKRLYEKAAVLDGSLKMIHTPPADLDVTNI---LQADEPTLTNTLDLSVLGK 114
 QY 115 DYGLAKDVIINANPASPPLSLVLRLLCERYVLTSTVTHSSVKNVPENLVKCFG-EOA 173
 DB 115 DLGALRDVIINANPASPPLSLVLRLLCERYVLTSTVTHSSVKNVPENLVKCFG-EOA 174
 QY 174 RKQSHHEVQLGFTLLWKAVNPKTQMKFSVQTMCPISGEENIARFLSLRG-QKNAVTLTL 232
 DB 175 HSYAHRRFQLGFTLLWKAVNPKTQMKFSVQTMCPISGEENIARFLSLRG-QKNAVTLTL 234
 QY 233 IDSWDIIMFOLREGSKKKAIVFRSMNSALGRSPMLVGNELTVADVLTMSVLQQTGSS 292
 DB 235 MDGWDVTLFOLABEGSKKKAIVFRSMNSALGRSPMLVGNELTVADVLTMSVLQQTGSS 294
 QY 293 GAAPTNVORWMLKSCENLAPFSTALDLK 320
 DB 295 -SAPANVORWMLKSCENLAPFSTALDLK 321

RESULT 11
 ID 0707A3 PRELIMINARY; PRT; 340 AA.
 AC 0707A3;
 DT 01-MAR-2004 (Tremblrel. 26, Created)
 DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)
 DT AGC5808 (Fragment).
 GN Name=agCG50514; ORFNames=ENSGANG0000011827;
 OS Anopheles gambiae str. PEST.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
 OX NCBI_TaxID=180454;
 RP SEQUENCE FROM N.A.

RC STRAIN=PEST;
 RA Anopheles Genome Sequencing Consortium;
 RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -! CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL, AAB01008960; EAA10870.1; -
 FT NON TER 1
 SQ SEQUENCE 340 AA; 37467 MW; FDD08525CD4EB976 CRC64;

Query Match 14.9%; Score 247; DB 2; Length 340;
 Best Local Similarity 28.7%; Pred. No. 6.1e-12;
 Matches 94; Conservative 49; Mismatches 118; Indels 66; Gaps 13;

QY 3 MYGVKRYHGSAPLRLVELPTCMYRLPNVHASKTSP-----ATD---AGHY----- 44
 DB 12 MYRLKPVN--NODVCELPCTCMYRLPYCAVSHDKRASADSATGLAANGPDPMSIED 69
 QY 45 ---QETSPPSLOAESRODDILKRLYEKAAVLDGSLKMIHTPPADLDV-TNLLQADEPT 100
 DB 70 NLLQKAKNDELRLMAERQORVLAQLAEK-----KEIMARTELKLNANPVAQPST 122
 QY 101 -----LATNTLDLSVLGDKDYALKDVIINANPASPPLSLVLRLLCERYVLTSTVH 153
 DB 123 PLKSKRAQKAEIRINLT-----CLQDFVNVASPEYVPSILALKNMLKRLNLQVECF 174
 QY 154 THSSVKNVPE-----NLVKCGEQAQRKSRHEVGLFTLWKV--PKTQMKFSVQTMCP 206
 DB 175 THSTVPKLSSEALAFQNAVTAAGTAANLPR---IKVTLKKNVAGAYTEMITSPTSYP 230
 QY 207 IEGEGNIARFLPSL-----FQKKNNAVTLTLDISWDIIMFOLREGSKKKAIVFRSMN 260
 DB 221 ICGEVNIIARVLSRCGPSEFNEQDN---VDSVDIILACVYLLKKNVKKQQLIRTLG 287
 QY 261 SALGRSPMLVGNELTVADVLTMSVLQ 287
 DB 288 AKLGKRAAGFGGADLSLCDIAFTSAVKQ 314

RESULT 12
 ID 07KUM5 PRELIMINARY; PRT; 301 AA.
 AC 07KUM5;
 DT 05-JUL-2004 (Tremblrel. 27, Created)
 DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
 DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
 DE CG12304-PB.
 GN ORFNames=CG12304;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=201956006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazer R.G., Chapple M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abril J.F., Abmayant A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Bendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu B., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dong U.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.U., Evariolista C.C., Ferraz C., Fertiera S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

| QY | 3 | MYGVKPEYHGGASAPLRYELPTCYMRLPNY-----HSKTSPTADGHNQETSEPSLOA | 54 |
|-----------------------|---|---|-----|
| Db | 1 | MYBLKTL---LQFOPIKILPTCYMRPLKXNVLASLADSLAGSSSTAS TSA--STSSCDDTAS | 54 |
| QY | 55 | LESRODDILIKRLYELKAAVDGISKMIHTPADLDPVTNIILOADEPTTLATNTLIDNSVLGK | 114 |
| Db | 55 | VAARQEKVKLQLEELKAOLGQIR-----AGLGVCG--KTFQHTTAFOG-----GL | 98 |
| QY | 115 | DYGALKDVIYNANPASPPLSLIVLHRLLCERYRVLSTVHTHSSVKNV-----PENLVK | 167 |
| Db | 99 | KEVPLDDVVINHPNPFIFUYALLAKNMRNLVYTDVTFPHSTMAIDIPGAARFEANLAK | 15 |
| QY | 168 | CFGEQARKOSRHEHYLGFTLLIMKXVPTKOMKFSVOTCPRIEKGNTARFLFSLFGQKHA | 22 |
| Db | 159 | VEPNPALP-----KISTVLLMKRCHEMTMISSTPMVPIFYGEVNIIRKILRGVPAEYX | 21 |
| QY | 228 | VTLTL---IDSKVDIAMEFQIRESSKEKAAVFSNMSALGRSPMLVGNELTVADVLMV | 28 |
| Db | 213 | BGSPPLNCNEIDVLDCYOLLRCNTHKTVAMVRLDKRLQKQYFGGSGMSVADVGYS | 27 |
| QY | 285 | L 285 | |
| Db | 273 | L 273 | |
| RESULT 13 | | | |
| Q8T060 | ID | PRELIMINARY; PRT; 313 AA. | |
| Q8T060 | AC | Q8T060; | |
| DT | 01-JUN-2002 | (TREMBLrel. 21, Created) | |
| DT | 01-MAR-2004 | (TREMBLrel. 26, Last sequence update) | |
| DT | 01-MAR-2004 | (TREMBLrel. 26, Last annotation update) | |
| DE | LD25772P | (Fragment). | |
| GN | ORFNames=CG12304; | | |
| OS | Drosophila melanogaster (Fruit fly). | | |
| OC | Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; | | |
| OC | Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; | | |
| OC | Ephydroidea; Drosophilidae; Drosophila. | | |
| OX | NCBI_TaxID=7227; | | |
| RN | (1) | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN=Berkeley; | | |
| RA | Stapleton M., Broketein P., Hong L., Agbayani A., Carlson J., | | |
| RA | Champe W., Chavez C., Dorsett V., Drensek D., Fattán D., Frise E., | | |
| RA | George R., Gonzalez M., Guatin H., Krommiller B., Li P., Liao G., | | |
| RA | Miranda A., Mungall C.J., Nunoo J., Paclob J., Paragas V., Park S., | | |
| RA | Patel S., Phuanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M., | | |
| RA | Celniker S.; | | |
| RL | Submitted (DRC-2003) to the EMBL/GenBank/DBJ databases. | | |
| DR | EMBL; AY069537; ALJ39682.2; - | | |
| DR | FlyBase; FBgn0036515; CG12304. | | |
| DR | InterPro; IPR004046; GST_Cterm. | | |
| DR | InterPro; IPR010987; GST_C_1like. | | |
| DR | Pfam; PF00043; GST_C; 1. | | |
| FT | NOM TER | | |
| SEQ | SEQUENCE | 313 AA; 34522 MW; 5AF9C801877F8531 CRC64; | |
| Query Match | 13.5%; | Score 224; DB 2; Length 313; | |
| Best Local Similarity | 28.6%; | Pred. No. 4.3e-10; | |
| Matches | 86; | Conservative 47; Mismatches 122; Indels 46; Gaps 9 | |
| QY | 3 | MYGVKPEYHGGASAPLRYELPTCYMRLPNY-----HSKTSPTADGHNQETSEPSLOA | 54 |
| Db | 13 | MYBLKTL---LQFOPIKILPTCYMRPLKXNVLASLADSLAGSSSTAS TSA--STSSCDDTAS | 66 |
| QY | 55 | LESRODDILIKRLYELKAAVDGISKMIHTPADLDPVTNIILOADEPTTLATNTLIDNSVLGK | 114 |
| Db | 67 | VAARQEKVKLQLEELKAOLGQIR-----AGLGVCG--KTFQHTTAFOG-----GL | 110 |
| QY | 115 | DYGALKDVIYNANPASPPLSLIVLHRLLCERYRVLSTVHTHSSVKNV-----PENLVK | 167 |
| Db | 111 | KEVPLDDVVINHPNPFIFUYALLAKNMRNLVYTDVTFPHSTMAIDIPGAARFEANLAK | 170 |

```
QY 168 CFGEQARKOSRHEYOQLGFTLLIKWNPXTOMKESVQTMCEPIEGEGNIARELFSLFGQIKHNA 227
DB 171 VFNVPALP-----KISVTLIKWNCNHTMISSPTMYVPIYGEVINIRILGVAPEXY 224
QY 228 VTLTL---IDSWVDIAMFOLREGSSKEKAIVFRSNMNSALGRSPWLVEGNETLVADVVLMSV 284
DB 225 EGSPICNEIDVLIDICYQLLRKNTHTKYAMVRLDLKRLKOQYRFGSGMSVADYGVSS 284
QY 285 L 285
DB 285 L 285

RESULT 14
GENKX4 PRELIMINARY; PRT; 322 AA.
ID 06NKM4
AC 06NKM4;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE LPI211AD (Fragment).
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Carlson J., Chavez C., Friese E., George R., Pacleb J.,
RA Park S., Wan K., Yu C., Rubin G.M., Ceiniker S.;
RU Submitted (May-2004) to the EMBL/Genbank/DBJ databases.
DR EMBL; BT012670; AAT08476.1; -.
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR010987; GST_C_1like.
DR Pfam; PF000043; GST_C_1.
FT NON TER
PT 322
SQ SEQUENCE 322 AA; 35528 MW; 82A15927DF928DAE CRC64;

Query Match 13.5%; Score 224; DB 2; Length 322;
Beet Local Similarity 28.6%; Pred. No. 4,5e-10;
Matches 86; Conservative 47; Mismatches 122; Indels 46; Gaps 9;

QY 3 MYQVPRYHGSGAFLRVELPTCMYRLPNV-----HSKTSPTATDAGHVQSTSEPSLQA 54
DB 22 MELTLTL---LPQFDIKLPTCMYRLPNVSLADSLASGSSASTA---STSSCDTAS 75
QY 55 LESRODDILKRLYEKKAADVGLSKMHTPDADLDVTNIIQADDEPTLATNTLDLSVLGK 114
DB 76 VAAREKVLKOLELKAQIQGR-----AGLGVC--KTFQHTTAPONG-----GL 119
QY 115 DYGALKDIYINANPASPSPLSLIVLHRLCERYRVUSTVTHSSVKNV-----PENLVK 167
DB 120 KEVLPQDVVINGHPNFIPIYALLALKNAMENLYTIDVKTFTHTMADIGPARAREFEANLAK 179
QY 168 CFGEQARKOSRHEYOQLGFTLLIKWNPXTOMKESVQTMCEPIEGEGNIARELFSLFGQIKHNA 227
DB 180 VFNVPALP-----KISVTLIKWNCNHTMISSPTMYVPIYGEVINIRILGVAPEXY 233
QY 228 VTLTL---IDSWVDIAMFOLREGSSKEKAIVFRSNMNSALGRSPWLVEGNETLVADVVLMSV 284
DB 234 EGSPICNEIDVLIDICYQLLRKNTHTKYAMVRLDLKRLKOQYRFGSGMSVADYGVSS 293
QY 285 L 285
DB 294 L 294

RESULT 15
MCA2_DROME STANDARD; PRT; 334 AA.
AC 09VUR3;
DT 16-OCT-2001 (rel. 40, Created)
```

```
DT 16-OCT-2001 (rel. 40, Last sequence update)
DT 25-JAN-2005 (rel. 46, Last annotation update)
DE Prohibitin multisynthetase complex auxiliary component p38.
CN ORFNames=CG12304;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoekins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklo G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup E.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glaeser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kamnitsos J.A., Ketchum K.A.,
RA Kimmel B.B., Kodira C.D., Kratz C., Kravitz S., Kulp D., Lai Z.,
RA Laske P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milhina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Ruskern D.R., Pacleb J.M.,
RA Palazolo M., Picman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitek R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RN [2]
RP GENOME REANNOTATION.
RX MEDLINE=24246069; PubMed=12537572;
RA Mista S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Batencourt B.R., Ceiniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review."
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
CC -!- FUNCTION: Probable core protein of the multisynthetase complex
CC that serves as a template for the assembly of the supramolecular
CC structure (By similarity).
CC -!- SUBUNIT: Component of the multisynthetase complex which is
CC comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the
CC monospecific isoleucyl, leucyl, glutamyl, methionyl, lysyl,
CC arginyl, and asparaginyl-CRNA synthetases as well as three auxiliary
CC proteins, p18, p48 and p43 (By similarity).
CC -!- SIMILARITY: Contains 1 GST-like domain.
```

```

-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.ebi.ch/announce/
CC      or send an email to license@ebi.ch).
CC      -----
DR      EMBL; AE003530; AAF49612.1; -.
DR      InFrac; Q9VUR3; -.
DR      Flybase; FBgn0036515; CG12304.
DR      InterPro; IPR010987; GST_C_like.
DR      InterPro; IPR004046; GST_Cterm.
DR      Pfam; PF00043; GST_C; 1.
DR      KW      Protein biosynthesis.
FT      DOMAIN 280
SQ      SEQUENCE 334 AA; 36933 MW; B68FD70AE62190F CRC64;

Query Match      13.1%; Score 216.5; DB 1; Length 334;
Best Local Similarity 26.0%; Pred. No. 2e-09;
Matches 86; Conservative 48; Mismatches 124; Indels 73; Gaps 9;

QY      3 MYQYKPYHGSAPLRVLPFCMTRLPNV-----HSKTTSPATDAG-----42
DB      1 MYELKTL--LPQFDILPTCMYPLKNVSLAADSLASGSSTASTSCKLEANRID 57

QY      43 -----HVQETSEPSLOAESRQDILKRYELKAVDGLSKMHTPD 84
DB      58 RTGNMATCALDLSLGRQIQRLIKDDPTASVAPRQEKVLRKQLELKAQLGQIR-----110

QY      85 ADDDVYNILOADEPTLATNTLDLNSVLRGKDYALKDVIYINANPASPLSLVLRHLCE 144
DB      111 AGGVCG--KTFQHTTAFQNG-----GLKEVPLQDVVINGHPNFIYALLALKNAMRN 161

QY      145 RYRVLSTVHTSSVKNV-----PENTLYKCFGEQARKSRHEYOUGFTLIMKNVFKTOM 197
DB      162 LYTIIVKTFTHSTMAIDIGPARAREFANLAKVPVNPALP-----KISVTLIMKNCETEM 215

QY      198 KFSVQTCPIEGGNTARLFSLFGQKNAVTLTL--IDSWVDIAMFQLRGSSSEKXA 254
DB      216 ISSPTWVVPYIGEVNIIRYLRGVPAEYRYESGPLCNEIDVLDICYQLRCNTHKTQYA 275

QY      255 VFRSNMSALGRSPMLVGNELTVADVVLMSYL 285
DB      276 MVRLLDKRLQKQYFGSQMSVADGVYSSSL 306

```

Search completed: February 23, 2005, 13:57:14
 Job time : 62.9811 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:25:39 ; Search time 70.371 Seconds
(without alignment)
1758.725 Million cell updates/sec

Title: US-10-622-817-6
Perfect score: 1655
Sequence: 1 MPMYGVKRYHGGASAPLRYEL.....RWLKSCENLAPRSTALQLLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues
Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneeqp19808:*
2: geneeqp19908:*
3: geneeqp20005:*
4: geneeqp20015:*
5: geneeqp20025:*
6: geneeqp20035:*
7: geneeqp20035:*
8: geneeqp20045:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 1464 | 88.5 | 320 | 7 | AD58613 Human Pro |
| 2 | 1464 | 88.5 | 320 | 8 | ABM80843 Tumour-as |
| 3 | 1436 | 86.8 | 312 | 2 | AAW25776 TVL1 prot |
| 4 | 1436 | 86.8 | 312 | 8 | ADR86551 1-312 am |
| 5 | 1232 | 74.4 | 272 | 7 | ADC10204 Human NOV |
| 6 | 1069 | 64.6 | 229 | 8 | ADR86553 84-312 am |
| 7 | 735 | 44.4 | 161 | 8 | ADR86552 1-161 am |
| 8 | 248 | 15.0 | 51 | 5 | ABP01502 Human ORF |
| 9 | 216.5 | 13.1 | 334 | 4 | ABB62468 Drosophila |
| 10 | 148 | 8.9 | 263 | 4 | ABG23964 Novel hum |
| 11 | 126.5 | 7.6 | 716 | 3 | AAQ1411 Arabidops |
| 12 | 126.5 | 7.6 | 719 | 3 | AAQ1410 Arabidops |
| 13 | 126.5 | 7.6 | 748 | 3 | AAQ1409 Arabidops |
| 14 | 106 | 6.4 | 1512 | 4 | AAU04349 Mammalian |
| 15 | 106 | 6.4 | 1512 | 4 | AAW78732 Human pro |
| 16 | 106 | 6.4 | 1550 | 4 | AAW79716 Human pro |
| 17 | 104 | 6.3 | 980 | 8 | ADN23683 Bacterial |
| 18 | 102 | 6.2 | 1049 | 6 | ABM67090 Photocorhab |
| 19 | 101.5 | 6.1 | 724 | 4 | AAQ85046 Shrimp wh |
| 20 | 99 | 6.0 | 691 | 6 | ABU16445 Protein e |
| 21 | 98 | 5.9 | 243 | 4 | AAU58158 Proportionb |
| 22 | 98 | 5.9 | 243 | 6 | ABM54677 Proportionb |
| 23 | 98 | 5.9 | 644 | 6 | ABJ18952 Pathogen |
| 24 | 98 | 5.9 | 690 | 4 | AAU34345 Staphyloc |
| 25 | 98 | 5.9 | 690 | 6 | ABW1749 Staphyloc |

| | | | | | |
|----|------|-----|------|---|--------------------|
| 26 | 98 | 5.9 | 985 | 8 | ADN88622 Rat epid1 |
| 27 | 98 | 5.9 | 997 | 8 | ADN88620 Rat epid1 |
| 28 | 98 | 5.9 | 1004 | 8 | ADN88621 Rat epid1 |
| 29 | 98 | 5.9 | 1013 | 8 | ADN88594 Rat epid1 |
| 30 | 98 | 5.9 | 1016 | 8 | ADN88619 Rat epid1 |
| 31 | 98 | 5.9 | 1032 | 8 | ADN88618 Rat epid1 |
| 32 | 96.5 | 5.8 | 1398 | 8 | ABM83058 Human dia |
| 33 | 96.5 | 5.8 | 1398 | 8 | ABM83057 Human dia |
| 34 | 96.5 | 5.8 | 1440 | 8 | ADN03674 Herbicida |
| 35 | 96.5 | 5.8 | 1440 | 8 | ADP54124 Human PRO |
| 36 | 96.5 | 5.8 | 1440 | 8 | ADP23041 PRO polyp |
| 37 | 96 | 5.8 | 2273 | 2 | AAQ98811 Erythiphe |
| 38 | 95 | 5.7 | 722 | 8 | ADP20980 Bacterial |
| 39 | 94 | 5.7 | 436 | 5 | ABBS4068 Lactococc |
| 40 | 93.5 | 5.6 | 1049 | 5 | ABR91239 Herbicida |
| 41 | 93.5 | 5.6 | 1064 | 5 | AAE25162 RCH1.5 pr |
| 42 | 93.5 | 5.6 | 1090 | 4 | ABG17694 Novel hum |
| 43 | 93.5 | 5.6 | 1090 | 4 | ABG18331 Novel hum |
| 44 | 93 | 5.6 | 204 | 4 | ADS21772 Bacterial |
| 45 | 93 | 5.6 | 559 | 2 | AAI10682 Polymydr |

ALIGNMENTS

| | |
|--|--|
| RESULT 1 | |
| AD58613 | |
| ID AD58613 standard; protein; 320 AA. | |
| XX | |
| AC AD58613; | |
| XX | |
| DT 29-JAN-2004 (first entry) | |
| XX | |
| DE Human Protein Q13155, SEQ ID NO 4489. | |
| XX | |
| KW Human; pain; neuronal tissue; gene therapy; | |
| KW spinal segmental nerve injury; chronic constriction injury; CCI; | |
| KW spared nerve injury; SNR; Chung. | |
| XX | |
| OS Homo sapiens. | |
| XX | |
| PN WO2003016475-A2. | |
| XX | |
| PD 27-FEB-2003. | |
| XX | |
| PF 14-AUG-2002; 2002WO-US025765. | |
| XX | |
| PR 14-AUG-2001; 2001US-0312147P. | |
| PR 01-NOV-2001; 2001US-0346382P. | |
| XX | |
| PR 26-NOV-2001; 2001US-0333347P. | |
| XX | |
| PA (GENO) GEN HOSPITAL CORP. | |
| PA (FARB) BAYER AG. | |
| XX | |
| PI Woolf C, D'urso D, Befort K, Costigan M; | |
| XX | |
| DR WPI; 2003-268312/26. | |
| DR GENBANK; Q13155. | |
| XX | |
| PT New composition comprising two or more isolated polypeptides, useful for | |
| PT preparing a medicament for treating pain in an animal. | |
| XX | |
| PS Claim 1; Page; 1017pp; English. | |
| XX | |
| XX | |
| CC The invention discloses a composition comprising two or more isolated rat | |
| CC or human polynucleotides or a polynucleotide which represents a fragment, | |
| CC derivative or allelic variation of the nucleic acid sequence. Also | |
| CC claimed are a vector comprising the novel polynucleotide, a host cell | |
| CC comprising the vector, a method for identifying a nucleotide sequence | |
| CC which is differentially regulated in an animal subjected to pain and a | |
| CC kit to perform the method, an array, a method for identifying an agent | |
| CC that increases or decreases the expression of the polynucleotide sequence | |
| CC that is differentially expressed in neuronal tissue of a first animal | |

CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNII), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 320 AA;

Query Match 88.5%; Score 1464; DB 7; Length 320;
Best Local Similarity 87.2%; Pred. No. 7, 5e-142;
Matches 279; Conservative 16; Mismatches 25; Indels 0; Gaps 0;

QY 1 MPWYQVKPYHGGGAPRLVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRD 60
DB 1 MPWYQVKPYHGGGAPRLVELPTCMYRLPNVHGRSGYPAGAGHVQESNLSLQALESRD 60
QY 61 DILKRLYEELKAADVGSKMIHTPDADLVNTNIIQADEPTTLATNTLDNSVLGKDYGALK 120
DB 61 DILKRLYEELKAADVGSKMIHTPDADLVNTNIIQADEPTTLATNTLDNSVLGKDYGALK 120
QY 121 DIVINANPASPPSLVLHRLCERYVSLVTHSSVKNVBNELVKCGEQARQSRHE 180
DB 121 DIVINANPASPPSLVLHRLCERYVSLVTHSSVKNVBNELVKCGEQARQSRHE 180
QY 181 YOLGFTLLWKVNPKTQMKFSVQTMCPIDEGGNIAARLFSLFGQKHNAAVTLTLDMSWVIA 240
DB 181 YOLGFTLLWKVNPKTQMKFSVQTMCPIDEGGNIAARLFSLFGQKHNAAVTLTLDMSWVIA 240
QY 241 MEOLREGSSKEKAAYFRSNMSALGSRPMLVGNELTVADVLMVSVLQOTGGSGGAAPTNAVQ 300
DB 241 IFOLREGSSKEKAAYFRSNMSALGSRPMLVGNELTVADVLMVSVLQOTGGSGVTPVNAVQ 300
QY 301 RMLKSCENLAPFTALQLLK 320
DB 301 RMRSCENLAPFTALKLK 320

RESULT 2

ABM80843
ID ABM80843 standard; protein; 320 AA.

AC ABM80843;

DT 18-NOV-2004 (first entry)

DE Tumour-associated antigenic target (TAT) polypeptide PROB1501, SEQ.2179.

KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic.

OS Homo sapiens.

XX WO2004030615-A2.

PN 15-APR-2004.

XX 29-SEP-2003; 2003MO-US028547.
PF 02-OCT-2002; 2002US-041971P.
XX (GENTH) GENENTECH INC.
PA Wu TD, Zhang Z, Zhou Y,
PI WPI; 2004-347921/32.
XX N-PSDB; ACN38579.
DR New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX Claim 12; SEQ ID NO 2179; 7273bp; English.
PS The invention relates to human tumour-associated antigenic target (TAT)
XX polypeptides, and their related nucleic acids. The TAT polypeptides are
XX overexpressed in cancer tissues compared to normal tissues, and may thus
XX serve as effective targets for the diagnosis and treatment of cancer in
XX mammals. The invention also relates to nucleic acid and polypeptide
XX sequences at least 80% identical to the TAT nucleic acid and polypeptide
XX polypeptides; expression vectors and host cells comprising a TAT nucleic
XX acid; an antibody specific for a TAT polypeptide; a peptide or organic
XX molecule which binds to a TAT polypeptide; fusion proteins comprising a
XX TAT polypeptide; and methods and compositions for the treatment or
XX diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
XX antibodies, antagonists, binding molecules and compositions are useful
XX for diagnosing or treating a cell proliferative disorder associated with
XX increased TAT expression, particularly cancers such as breast cancer,
XX colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
XX cancer, pancreatic cancer, cervical cancer, cancers of the central
XX nervous system, melanoma and leukaemia. TAT nucleic acids may further be
XX used as hybridisation probes, in chromosome and gene mapping, in
XX chromosome identification and in gene therapy. The present sequence
XX represents a TAT polypeptide of the invention
SQ Sequence 320 AA;

Query Match 88.5%; Score 1464; DB 8; Length 320;
Best Local Similarity 87.2%; Pred. No. 7, 5e-142;
Matches 279; Conservative 16; Mismatches 25; Indels 0; Gaps 0;

QY 1 MPWYQVKPYHGGGAPRLVELPTCMYRLPNVHSKTTSPATDAGHVQETSEPSLQALESRD 60
DB 1 MPWYQVKPYHGGGAPRLVELPTCMYRLPNVHGRSGYPAGAGHVQESNLSLQALESRD 60
QY 61 DILKRLYEELKAADVGSKMIHTPDADLVNTNIIQADEPTTLATNTLDNSVLGKDYGALK 120
DB 61 DILKRLYEELKAADVGSKMIHTPDADLVNTNIIQADEPTTLATNTLDNSVLGKDYGALK 120
QY 121 DIVINANPASPPSLVLHRLCERYVSLVTHSSVKNVBNELVKCGEQARQSRHE 180
DB 121 DIVINANPASPPSLVLHRLCERYVSLVTHSSVKNVBNELVKCGEQARQSRHE 180
QY 181 YOLGFTLLWKVNPKTQMKFSVQTMCPIDEGGNIAARLFSLFGQKHNAAVTLTLDMSWVIA 240
DB 181 YOLGFTLLWKVNPKTQMKFSVQTMCPIDEGGNIAARLFSLFGQKHNAAVTLTLDMSWVIA 240
QY 241 MEOLREGSSKEKAAYFRSNMSALGSRPMLVGNELTVADVLMVSVLQOTGGSGGAAPTNAVQ 300
DB 241 IFOLREGSSKEKAAYFRSNMSALGSRPMLVGNELTVADVLMVSVLQOTGGSGVTPVNAVQ 300
QY 301 RMLKSCENLAPFTALQLLK 320
DB 301 RMRSCENLAPFTALKLK 320

RESULT 3

AAW25776

ID AAW25776 standard; protein; 312 AA.
 AC AAW25776;
 XX
 DT 19-DEC-1997 (first entry)
 XX
 DE JTV1 protein.
 XX
 KW JTV1; hPMS2; probe; detection; chromosome 7; deletion;
 KW mismatch repair gene; hereditary non-polyposis colorectal cancer;
 KW homologous recombination.
 XX
 OS Homo sapiens.
 XX
 PN MO9708312-A1.
 XX
 PD 06-MAR-1997.
 XX
 PF 26-AUG-1996; 96WO-US013598.
 XX
 PR 24-AUG-1995; 95US-00518862.
 XX
 PA (UYJO) UNIV JOHNS HOPKINS.
 XX
 PI Vogelerstein B, Kinzler KW, Nicolaides NC;
 XX
 DR WPI; 1997-179269/16.
 DR N-PSDB; AAT86182.
 XX
 PT Novel chromosome 7 gene, JTV1 - used for detecting chromosome 7
 PT deletions, and PMS2 promoter activity.
 XX
 PS Claim 5; Fig 2; 31pp; English.
 XX
 CC This sequence is JTV1 protein and is encoded by DNA isolated from human
 CC chromosome 7. The JTV1 coding sequence is located upstream from hPMS2.
 CC JTV1 cDNA can be used as probes to detect chromosome 7 deletions
 CC involving JTV1. Due to the overlapping promoter regions, deletions of
 CC JTV1 would also affect PMS2 (a mismatch repair gene) expression, leading
 CC to hereditary non-polyposis colorectal cancer. JTV1 can also be used to
 CC assay activity or competence of the PMS2 promoter region, the presence of
 CC JTV1 suggesting that the PMS2 promoter is intact. JTV1 sequences can also
 CC be used to guide homologous recombination at the PMS2 locus
 CC
 XX
 SQ Sequence 312 AA;
 Query Match 86.8%; Score 1436; DB 2; Length 312;
 Best Local Similarity 87.5%; Pred. No. 5,6e-139;
 Matches 273; Conservative 14; Mismatches 25; Indels 0; Gaps 0;
 QY 1 MPMYQVKRYHGGASAPLRLVLPCTCMYRLPNVHNSKTTSPATDAGHVQETSEPSIQALESRD 60
 DB 1 MPMYQVKRYHGGASAPLRLVLPCTCMYRLPNVHNSKTTSPATDAGHVQETSEPSIQALESRD 60
 QY 61 DILKRLYLEKAAVDSLSKMIHTPDADLVNTILQADEPTTLATNTLDLSNVGKDYGALK 120
 DB 61 DILKRLYLEKAAVDSLSKMIHTPDADLVNTILQADEPTTLATNTLDLSNVGKDYGALK 120
 QY 121 DIVINANPASPPLSLVLRLLCERYRVLSVTHSSVKNVBNLVKCFGEQARKQSRHE 180
 DB 121 DIVINANPASPPLSLVLRLLCERYRVLSVTHSSVKNVBNLVKCFGEQARKQSRHE 180
 QY 181 YQLGTTILWKNVPTQMKFSVOTMCPICEGGINAFRLSLFQCKKNATVLTLLIDSWDIA 240
 DB 181 YQLGTTILWKNVPTQMKFSVOTMCPICEGGINAFRLSLFQCKKNATVLTLLIDSWDIA 240
 QY 241 MFOLEGGSSKEKAAVFRSNNSALGRSPMLVGNELTVADVLSVYLQOTGGSSGAAPTNYQ 300
 DB 241 IFOLGEGSSKEKAAVFRSNNSALGRSPMLVGNELTVADVLSVYLQOTGGSSGAAPTNYQ 300
 QY 301 RMLKSCENLAMP 312
 DB 301 RMLKSCENLAMP 312

RESULT 4
 ID ADR86551 standard; protein; 312 AA.
 AC ADR86551;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE 1-312 amino acid sequence of p38/JTV-1 protein.
 XX
 KW p38/JTV-1; cytosstatic; cancer; leukemia; anticancer.
 XX
 OS Homo sapiens.
 XX
 PN EP1454628-A2.
 XX
 PD 08-SEP-2004.
 XX
 PF 09-SEP-2003; 2003EP-00020344.
 XX
 PR 03-MAR-2003; 2003KR-00013058.
 XX
 PA (UYSE-) UNIV SEOUL NAT IND FOUND.
 XX
 PI Kim S, Park B;
 XX
 DR WPI; 2004-627822/61.
 DR N-PSDB; ADR86548.
 XX
 PT New isolated p38/JTV-1 protein, useful as medicament for treating cancer
 PT e.g., stomach, liver, blood, bone, pancreatic, skin, head or neck cancer
 PT and cutaneous or intraocular melanoma, as well as for screening new
 PT anticancer agents.
 XX
 PS Claim 5; SEQ ID NO 4; 47pp; English.
 XX
 CC The present invention relates to an isolated p38/JTV-1 protein for use as
 CC medicament. The p38/JTV-1 protein or the pharmaceutical composition is
 CC useful as medicament for treating breast cancer, large intestinal cancer,
 CC lung cancer, small cell lung cancer, stomach cancer, liver cancer, blood
 CC cancer, bone cancer, pancreatic cancer, skin cancer, head or neck cancer,
 CC cutaneous or intraocular melanoma, uterine sarcoma, ovarian cancer,
 CC rectal cancer, anal cancer, colon cancer, fallopian tube carcinoma,
 CC endometrial carcinoma, cervical cancer, vulval cancer, vaginal carcinoma,
 CC Hodgkin's disease, esophageal cancer, small intestine cancer, endocrine
 CC cancer, thyroid cancer, parathyroid cancer, adrenal cancer, soft tissue
 CC tumour, urethral cancer, penile cancer, prostate cancer, chronic or acute
 CC leukemia, lymphocytic lymphoma, bladder cancer, kidney cancer, ureter
 CC cancer, renal cell carcinoma, renal pelvic carcinoma, CNS tumour, primary
 CC CNS lymphoma, bone marrow tumour, brain stem nerve gliomas, pituitary
 CC adenoma, or their combination. The protein is useful as a target for
 CC screening new anticancer agents. The present sequence represents the 1-
 CC 312 amino acid sequence of p38/JTV-1 protein.
 CC
 XX
 SQ Sequence 312 AA;
 Query Match 86.8%; Score 1436; DB 8; Length 312;
 Best Local Similarity 87.5%; Pred. No. 5,6e-139;
 Matches 273; Conservative 14; Mismatches 25; Indels 0; Gaps 0;
 QY 1 MPMYQVKRYHGGASAPLRLVLPCTCMYRLPNVHNSKTTSPATDAGHVQETSEPSIQALESRD 60
 DB 1 MPMYQVKRYHGGASAPLRLVLPCTCMYRLPNVHNSKTTSPATDAGHVQETSEPSIQALESRD 60
 QY 61 DILKRLYLEKAAVDSLSKMIHTPDADLVNTILQADEPTTLATNTLDLSNVGKDYGALK 120
 DB 61 DILKRLYLEKAAVDSLSKMIHTPDADLVNTILQADEPTTLATNTLDLSNVGKDYGALK 120
 QY 121 DIVINANPASPPLSLVLRLLCERYRVLSVTHSSVKNVBNLVKCFGEQARKQSRHE 180
 DB 121 DIVINANPASPPLSLVLRLLCERYRVLSVTHSSVKNVBNLVKCFGEQARKQSRHE 180

QY 181 YOLGFTLIMKVPKTMKFSVOTMCPFBEGNIAFLFSLFGOKHNAVTLTLINSWDIA 240
DB 181 YOLGFTLIMKVPKTMKFSVOTMCPFBEGNIAFLFSLFGOKHNAVTLTLINSWDIA 240
QY 241 MFOLBEGSSKEKAARFRSMNSALGRSPWLVGNETLVADVLMVLSVLTQOTGSSGAAPTNNQ 300
DB 241 IFOLKEGSSKEKAARFRSMNSALGRSPWLVGNETLVADVLMVLSVLTQOTGSSGAAPTNNQ 300
QY 301 RMLKSCENLAPF 312
DB 301 RMLKSCENLAPF 312

RESULT 5
ADCl0204
ID ADCl0204 standard; protein; 272 AA.

AC ADCl0204;
XX
DT 18-DEC-2003 (first entry)

XX Human NOVX polypeptide SEQ ID NO: 226.

XX cytostatic; antidiabetic; anorectic; cerebroprotective; neuroprotective;
XX antiinflammatory; gene therapy; antisense therapy; thymomimetic; NOVX;
XX pathology; cancer; diabetes; obesity; endocrine disorder; CNS disorder;
XX inflammatory disorder; chromosome mapping; tissue typing;
XX predictive medicine.

XX Homo sapiens.

XX WO2003000842-A2.

XX PD 03-JAN-2003.

XX PF 04-JUN-2002; 2002WO-US017443.

XX 04-JUN-2001; 2001US-0295607P.
PR 04-JUN-2001; 2001US-0295610P.
PR 06-JUN-2001; 2001US-0296404P.
PR 06-JUN-2001; 2001US-0296418P.
PR 07-JUN-2001; 2001US-0296575P.
PR 11-JUN-2001; 2001US-0297414P.
PR 12-JUN-2001; 2001US-0295573P.
PR 12-JUN-2001; 2001US-0297567P.
PR 14-JUN-2001; 2001US-0296285P.
PR 15-JUN-2001; 2001US-0296285P.
PR 18-JUN-2001; 2001US-0299133P.
PR 19-JUN-2001; 2001US-0299230P.
PR 21-JUN-2001; 2001US-0299949P.
PR 22-JUN-2001; 2001US-0300177P.
PR 26-JUN-2001; 2001US-0300883P.
PR 28-JUN-2001; 2001US-0301530P.
PR 28-JUN-2001; 2001US-0301530P.
PR 03-JUL-2001; 2001US-0302951P.
PR 31-JUL-2001; 2001US-0308890P.
PR 14-SEP-2001; 2001US-0322297P.
PR 25-SEP-2001; 2001US-0324659P.
PR 03-DEC-2001; 2001US-0337477P.
PR 14-DEC-2001; 2001US-0341562P.
PR 21-FEB-2002; 2002US-0358666P.
PR 21-FEB-2002; 2002US-0359122P.
PR 22-FEB-2002; 2002US-0359789P.
PR 22-FEB-2002; 2002US-0359034P.
PR 22-FEB-2002; 2002US-0359035P.
PR 22-FEB-2002; 2002US-0359121P.
PR 27-FEB-2002; 2002US-0359964P.
PR 01-MAR-2002; 2002US-0360858P.
PR 12-MAR-2002; 2002US-0363430P.
PR 12-MAR-2002; 2002US-0363676P.
PR 10-APR-2002; 2002US-0371346P.
PR 10-MAY-2002; 2002US-0379444P.

PR 04-JUN-2002; 2002US-00379444.
XX (CURA-) CURAGEN CORP.
XX

PI Agee ML, Anderson DM, Berghs C, Casman SJ, Catterton E;
PI Dipippo VA, Edinger SR, Eissen A, Ellerman K, Gangolli EA;
PI Gerlach VL, Gorman L, Guo X, Hermann JJ, Hjalte T, Ji W, Kekuda R;
PI Khramtsov NV, Li L, Liu X, Malyanar UM, Miller CE, Miller I;
PI Ort T, Padigaru M, Paturajan M, Pena CE, Raetelli L, Rieger DK;
PI Rothenberg ME, Shenoy SG, Shinkets RA, Smithson G, Spaderna SK;
PI Spytek KA, Stone DJ, Vernet CM, Zhong H, Zhong M, Alsebrook JP;
PI Burgess CE, Lepley DM;

XX WPI; 2003-210149/20.
DR N-PSDB; ADCl0203.

PT New isolated NOVX polypeptides and nucleic acid molecules useful for
PT treating, preventing and diagnosing pathological conditions with NOVX-
PT associated disorders, such as cancer, obesity, diabetes and inflammatory
PT or CNS diseases.

XX Claim 1; SEQ ID NO 226; 772pp; English.

XX The invention relates to novel isolated polypeptides, mature form of the
XX polypeptide, a sequence that is 95% identical to the polypeptide or the
XX polypeptide comprising one or more conservative substitutions. The NOVX
XX polypeptide is useful for treating or preventing a pathology associated
XX with the polypeptide e.g. disorders associated with aberrant expression
XX or activity of the polypeptide, such as cancer, diabetes, obesity, and
XX endocrine, CNS and inflammatory disorders. They can also be used in
XX various detection and screening assays, chromosome mapping, tissue typing
XX and predictive medicine. This sequence corresponds to one of the
XX polypeptides of the invention.

XX Sequence 272 AA;

Query Match 74.4%; Score 1232; DB 7; Length 272;

Best Local Similarity 88.4%; Pred. No. 5.1e-118;
Matches 237; Conservative 12; Mismatches 19; Indels 0; Gaps 0;

QY 45 QETSEBSLOLESRODDILKRLYEKAAVAGLSKMIHTPDADLDVNTILQADEPTTLATN 104
DB 5 QESNSLSLOLESRODDILKRLYEKAAVAGLSKMIHTPDADLDVNTILQADEPTTLATN 64

QY 105 TLDINSVLGKDYALDVIYNNPASPPLSLVYHRLCERRVYSTYHTSSVKNVBN 164
DB 65 ALDINSVLGKDYALDVIYNNPASPPLSLVYHRLCERRVYSTYHTSSVKNVBN 124

QY 165 LVKCFGEQARKOSRHEVQLGFTLIMKVPKTMKFSVOTMCPFBEGNIAFLFSLFGOK 224
DB 125 LKCFGEQARKOPRODYGFTLIMKVPKTMKFSVOTMCPFBEGNIAFLFSLFGOK 184

QY 225 HNAVVTLLIDSWDIAFMQLBEGSSKEKAARFRSMNSALGRSPWLVGNETLVADVLMVSV 284
DB 185 HNAVVTLLIDSWDIAFMQLBEGSSKEKAARFRSMNSALGRSPWLVGNETLVADVLMVSV 244

QY 285 LQOTGSSGAAPTNNQRMMLKSCENLAPF 312
DB 245 LQOTGSSGAAPTNNQRMMLKSCENLAPF 272

RESULT 6
AD86553
ID AD86553 standard; protein; 229 AA.

XX AD86553;

DT 18-NOV-2004 (first entry)

DE 84-312 amino acid sequence of p38/JTV-1 protein.

XX p38/JTV-1; cytostatic; cancer; leukemia; anticancer.

OS Homo sapiens.
XX
XX EPI454628-A2.
XX
PD 08-SEP-2004.
XX
XX 09-SEP-2003; 2003EP-00020344.
XX
XX 03-MAR-2003; 2003KR-00013058.
XX
XX (UYSE-) UNIV SEOUL NAT IND FOUND.
PA
XX Kim S, Park B;
XX
XX WPI; 2004-627822/61.
DR N-PSDB; ADR86550.
XX
XX New isolated p38/JTV-1 protein, useful as medicament for treating cancer
PT e.g., stomach, liver, blood, bone, pancreatic, skin, head or neck cancer
PT and cutaneous or intraocular melanoma, as well as for screening new
PT anticancer agents.
XX
XX Claim 5; SEQ ID NO 6; 47bp; English.
XX
XX The present invention relates to an isolated p38/JTV-1 protein for use as
CC medicament. The p38/JTV-1 protein or the pharmaceutical composition is
CC useful as medicament for treating breast cancer, large intestinal cancer,
CC lung cancer, small cell lung cancer, stomach cancer, liver cancer, blood
CC cancer, bone cancer, pancreatic cancer, skin cancer, head or neck cancer,
CC cutaneous or intraocular melanoma, uterine sarcoma, ovarian cancer,
CC rectal cancer, anal cancer, colon cancer, fallopian tube carcinoma,
CC endometrial carcinoma, cervical cancer, vulval cancer, vaginal carcinoma,
CC Hodgkin's disease, esophageal cancer, small intestine cancer, endocrine
CC cancer, thyroid cancer, parathyroid cancer, adrenal cancer, soft tissue
CC tumour, urethral cancer, penile cancer, prostate cancer, chronic or acute
CC leukemia, lymphocytic lymphoma, bladder cancer, kidney cancer, ureter
CC cancer, renal cell carcinoma, renal pelvic carcinoma, CNS tumour, primary
CC CNS lymphoma, bone marrow tumour, brain stem nerve gliomas, pituitary
CC adenoma, or their combination. The protein is useful as a target for
CC screening new anticancer agents. The present sequence represents the 84-
CC 312 amino acid sequence of p38/JTV-1 protein.
XX
XX Sequence 229 AA;
SQ
Query Match 64.6%; Score 1069; DB 8; Length 229;
Best Local Similarity 88.2%; Pred. No. 2,6e-101;
Matches 202; Conservative 12; Mismatches 15; Indels 0; Gaps 0;
QY 84 DADLDVTNIIQADEPTTLATNTLDNSVVGKDYGALKDVIINANPASPPLSLTVLHRLLC 143
DB 1 DADLDVTNIIQADEPTTLATNTLDNSVVGKDYGALKDVIINANPASPPLSLTVLHRLLC 60
QY 144 ERYRLSTVHTHSVSKVNPENLVKCFGEQARKQSBHEYOLGFTLIWKVNPKTOMKFSVOT 203
DB 61 EHFRLSTVHTHSVSKVNPENLVKCFGEQARKQSBHEYOLGFTLIWKVNPKTOMKFSVOT 120
QY 204 MCPRIEGEGNIAFLSLRGQKKNATVLTLLDSWVDIAMQRLREGSSKEGAAYFRSNNSL 263
DB 121 MCPRIEGEGNIAFLSLRGQKKNATVLTLLDSWVDIAMQRLREGSSKEGAAYFRSNNSL 180
QY 264 GRSPLVAGNELTVADVLMVSVLQOTGSSGGAFTVVRKMLKSCENLAFF 312
DB 161 GRSPLVAGNELTVADVLMVSVLQOTGSSGGAFTVVRKMLKSCENLAFF 229

RESULT 7
ADR86552
ID ADR86552 standard; protein; 161 AA.
XX
XX ADR86552;
XX
XX 18-NOV-2004 (first entry)
XX

DE 1-161 amino acid sequence of p38/JTV-1 protein.
XX
XX p38/JTV-1; Cytostatic; cancer; leukemia; anticancer.
XX
XX Homo sapiens.
XX
XX EPI454628-A2.
XX
XX 08-SEP-2004.
XX
XX 09-SEP-2003; 2003EP-00020344.
XX
XX 03-MAR-2003; 2003KR-00013058.
XX
XX (UYSE-) UNIV SEOUL NAT IND FOUND.
PA
XX Kim S, Park B;
XX
XX WPI; 2004-627822/61.
DR N-PSDB; ADR86549.
XX
XX New isolated p38/JTV-1 protein, useful as medicament for treating cancer
PT e.g., stomach, liver, blood, bone, pancreatic, skin, head or neck cancer
PT and cutaneous or intraocular melanoma, as well as for screening new
PT anticancer agents.
XX
XX Claim 5; SEQ ID NO 5; 47bp; English.
XX
XX The present invention relates to an isolated p38/JTV-1 protein for use as
CC medicament. The p38/JTV-1 protein or the pharmaceutical composition is
CC useful as medicament for treating breast cancer, large intestinal cancer,
CC lung cancer, small cell lung cancer, stomach cancer, liver cancer, blood
CC cancer, bone cancer, pancreatic cancer, skin cancer, head or neck cancer,
CC cutaneous or intraocular melanoma, uterine sarcoma, ovarian cancer,
CC rectal cancer, anal cancer, colon cancer, fallopian tube carcinoma,
CC endometrial carcinoma, cervical cancer, vulval cancer, vaginal carcinoma,
CC Hodgkin's disease, esophageal cancer, small intestine cancer, endocrine
CC cancer, thyroid cancer, parathyroid cancer, adrenal cancer, chronic or acute
CC leukemia, lymphocytic lymphoma, bladder cancer, kidney cancer, ureter
CC cancer, renal cell carcinoma, renal pelvic carcinoma, CNS tumour, primary
CC CNS lymphoma, bone marrow tumour, brain stem nerve gliomas, pituitary
CC adenoma, or their combination. The protein is useful as a target for
CC screening new anticancer agents. The present sequence represents the 1-
CC 161 amino acid sequence of p38/JTV-1 protein.
XX
XX Sequence 161 AA;
SQ
Query Match 44.4%; Score 735; DB 8; Length 161;
Best Local Similarity 88.8%; Pred. No. 4.4e-67;
Matches 143; Conservative 5; Mismatches 13; Indels 0; Gaps 0;
QY 1 MEMVQVKEHYGGASAPLRYELPTCMYRLPNVHSKTSPTADAGHVBETSEPSIQALESROD 60
DB 1 MEMVQVKEHYGGASAPLRYELPTCMYRLPNVHSKTSPTADAGHVBETSEPSIQALESROD 60
QY 61 DILKRLYELKKAADVGLSKVITPPDADLVNTIIQADEPTTLATNTLDNSVVGKDYGALK 120
DB 61 DILKRLYELKKAADVGLSKVITPPDADLVNTIIQADEPTTLATNTLDNSVVGKDYGALK 120
QY 121 DIVINANPASPPLSLTVLHRLLCERVRLSTVHTSSVKNV 161
DB 121 DIVINANPASPPLSLTVLHRLLCERVRLSTVHTSSVKNV 161

RESULT 8
ABP01502
ID ABP01502 standard; protein; 51 AA.
XX
XX ABP01502;
XX
XX 24-JUN-2002 (first entry)
XX

DE Human ORFX protein sequence SEQ ID NO:2986.
XX
XX Human, open reading frame; ORFX, gene therapy; cancer; cirrhosis;
KM hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KM degenerative disorder; osteoarthritis; neurodegenerative disorder;
KM cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KM hyperension; hypothyroidism; cholesterol ester storage disease;
KM immune deficiency; immune disorder; infectious disease;
KM autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KM myasthenia gravis.
XX
XX Homo sapiens.
OS
XX MO200192523-A2.
PN
XX
XX 06-DEC-2001.
PD
XX
XX 29-MAY-2001; 2001WO-US010836.
PF
XX
XX 30-MAY-2000; 2000US-0206132P.
PR
XX 29-AUG-2000; 2000US-0228716P.
XX
XX (CUBA-) CUBAGEN CORP.
XX
XX Shinkets RA, Leach MD;
PI
XX WPI; 2002-106308/14.
DR
XX N-PSDB; ABN17254.
XX
XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders.
XX
XX Disclosure; SEQ ID NO 2986; 1037P; English.
PS
XX
XX The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, diabetes mellitus, systemic
CC transplantion, cardiovascular diseases, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune inflammatory eye disease, graft-versus-host
CC disease and autoimmune inflammatory diseases. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage. N.B. The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 51 AA;
SQ
Query Match 15.0%; Score 248; DB 5; Length 51;
Best Local Similarity 96.1%; Pred. No. 1.3e-11;
Matches 49; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 74 DGLSKMHTPPADLDVTNIIQADEPTTLATNTLDLNSVLGKDYALDVIY 124
DB 1 DGLSKMHTPPADLDVTNIIQADEPTTLATNTLDLNSVLGKDYALDVIY 51
RESULT 9
AB62468

ID ABB62468 standard; protein; 334 AA.
XX
XX ABB62468;
AC
XX 26-MAR-2002 (first entry)
DT
XX Drosophila melanogaster polypeptide SEQ ID NO 14196.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical.
XX
XX Drosophila melanogaster.
OS
XX
XX MO200171042-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US009231.
PF
XX
XX 23-MAR-2000; 2000US-0191637P.
PR
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li FWD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR
XX N-PSDB; ABL06571.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 14196; 21P + Sequence Listing; English.
PS
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins (ABBS7737-
CC ABBS2072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 334 AA;
SQ
Query Match 13.1%; Score 216.5; DB 4; Length 334;
Best Local Similarity 26.0%; Pred. No. 4.8e-13;
Matches 86; Conservative 48; Mismatches 124; Indels 73; Gaps 9;
OY 3 MYQVPRYHGSAPLREVELPTCMYRLPNV-----HSKTTSPATDAG----- 42
DB 1 MYELKTL---LPQFDIKLPTCMYRLPNVSLADSLASGSSYTSASTSSCKLEANRID 57
OY 43 -----HVGTSRPSLQALRESRDDILKRLYELKAAVVDGLSKMHTPD 84
DB 58 RTGRNAATCAALDLSLGRQIQRLKODTPASVAARBEKVKQLEIKAKDLGGIR----- 110
OY 85 ADLDVTNIIQADEPTTLATNTLDLNSVLGKDYALDVIYANPASPPLSLIVLRLLCE 144
DB 111 AGLGVCG--KTQHTTAPONG-----GLKEVPADVDVINCHPNFIPYALALAKNAWRN 161
OY 145 RYRVLSIVTHSSVKNV-----PENLVKCGEQARKQSRHRYOLGFTLIMKNVPTQM 197
DB 162 LVTIDVKTFTSTMDADIGPARPEFANLAKYVNPALP-----KISVTLWKCEHTEM 215
OY 198 KPSVQVMCIIEBEGNIAFLFSLFGQKNAVLTLT---IDSWVDIAMFQLRBGSSEKXA 254
DB 216 ISSPTMYVPIYGVNIRIRYIGRVGPAEYRYBSSPLCNELDVLDCYQLIRNTHKTQVA 275
OY 255 VFRSNNSALGRSPWLIVGNELTVADVVLMSVL 285

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:36:20 ; Search time 17.81 Seconds
(without alignments)
1341.256 Million cell updates/sec

Title: US-10-622-817-6

Perfect score: 1655

Sequence: 1 MPWQVQKHYHGSAFLRVEL.....RWLKSCEMLAPSTALQLLK 320

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PTCUTS.COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------------------|-------------------|
| 1 | 1464 | 86.5 | 341 | 4 US-09-949-016-11312 | Sequence 11312, A |
| 2 | 1436 | 86.8 | 312 | 2 US-08-518-862C-2 | Sequence 2, Appl |
| 3 | 106 | 6.4 | 1512 | 3 US-08-443-184-48 | Sequence 48, Appl |
| 4 | 96.5 | 5.8 | 1440 | 3 US-09-357-251-37 | Sequence 37, Appl |
| 5 | 90.5 | 5.5 | 359 | 3 US-09-540-824-2 | Sequence 2, Appl |
| 6 | 90 | 5.4 | 559 | 2 US-08-756-317-10 | Sequence 10, Appl |
| 7 | 90 | 5.4 | 559 | 4 US-09-091-609-4 | Sequence 4, Appl |
| 8 | 90 | 5.4 | 2954 | 4 US-09-150-867-1 | Sequence 1, Appl |
| 9 | 90 | 5.4 | 5215 | 3 US-09-105-537-2 | Sequence 2, Appl |
| 10 | 89 | 5.4 | 1056 | 4 US-09-595-684B-29 | Sequence 29, Appl |
| 11 | 89 | 5.4 | 1057 | 3 US-09-541-782-10 | Sequence 10, Appl |
| 12 | 89 | 5.4 | 1057 | 4 US-09-723-820-10 | Sequence 10, Appl |
| 13 | 89 | 5.4 | 1057 | 4 US-10-270-085-10 | Sequence 10, Appl |
| 14 | 89 | 5.4 | 1287 | 4 US-09-949-016-7826 | Sequence 7826, Ap |
| 15 | 88 | 5.3 | 712 | 4 US-09-489-039A-10736 | Sequence 10736, A |
| 16 | 88 | 5.3 | 778 | 4 US-09-583-110-3930 | Sequence 3930, Ap |
| 17 | 88 | 5.3 | 786 | 4 US-09-107-433-3893 | Sequence 3893, Ap |
| 18 | 86.5 | 5.2 | 520 | 4 US-09-248-796A-17644 | Sequence 17644, A |
| 19 | 86.5 | 5.2 | 967 | 4 US-09-540-236-2449 | Sequence 2449, Ap |
| 20 | 86 | 5.2 | 535 | 4 US-09-489-039A-11461 | Sequence 11461, A |
| 21 | 85.5 | 5.2 | 556 | 4 US-09-134-000C-6329 | Sequence 6329, Ap |
| 22 | 85.5 | 5.2 | 608 | 4 US-09-284-768A-4 | Sequence 4, Appl |
| 23 | 85.5 | 5.2 | 657 | 4 US-09-284-768A-7 | Sequence 7, Appl |
| 24 | 85 | 5.1 | 222 | 4 US-09-248-796A-18146 | Sequence 18146, A |
| 25 | 85 | 5.1 | 302 | 4 US-09-107-532A-6924 | Sequence 6924, Ap |
| 26 | 85 | 5.1 | 559 | 4 US-09-672-749-2 | Sequence 2, Appl |
| 27 | 85 | 5.1 | 1971 | 4 US-09-914-272A-1 | Sequence 1, Appl |

| | | | | | |
|----|------|-----|------|------------------------|-------------------|
| 28 | 85 | 5.1 | 1971 | 4 US-10-638-333-1 | Sequence 1, Appl |
| 29 | 84 | 5.1 | 443 | 4 US-09-949-016-10598 | Sequence 10598, A |
| 30 | 84 | 5.1 | 646 | 4 US-09-248-796A-18678 | Sequence 18678, A |
| 31 | 84 | 5.1 | 1057 | 4 US-09-428-156B-2 | Sequence 2, Appl |
| 32 | 83.5 | 5.0 | 360 | 4 US-09-949-016-10589 | Sequence 10589, A |
| 33 | 83.5 | 5.0 | 475 | 4 US-09-328-352-5379 | Sequence 5379, Ap |
| 34 | 83.5 | 5.0 | 519 | 4 US-09-489-039A-8208 | Sequence 8208, Ap |
| 35 | 83.5 | 5.0 | 515 | 4 US-09-543-681A-6505 | Sequence 6505, Ap |
| 36 | 83.5 | 5.0 | 789 | 3 US-09-002-285-84 | Sequence 84, Appl |
| 37 | 83.5 | 5.0 | 789 | 4 US-09-589-477-84 | Sequence 84, Appl |
| 38 | 83.5 | 5.0 | 789 | 4 US-10-099-285A-84 | Sequence 84, Appl |
| 39 | 83.5 | 5.0 | 959 | 4 US-09-543-681A-6879 | Sequence 6879, Ap |
| 40 | 83 | 5.0 | 434 | 3 US-09-012-072-4 | Sequence 4, Appl |
| 41 | 83 | 5.0 | 434 | 3 US-09-120-601-4 | Sequence 4, Appl |
| 42 | 83 | 5.0 | 448 | 3 US-09-120-601-6 | Sequence 6, Appl |
| 43 | 83 | 5.0 | 559 | 4 US-09-821-016-1 | Sequence 1, Appl |
| 44 | 83 | 5.0 | 559 | 4 US-10-266-787-1 | Sequence 1, Appl |
| 45 | 83 | 5.0 | 767 | 3 US-08-836-567-8 | Sequence 8, Appl |

ALIGNMENTS

```
RESULT 1
US-09-949-016-11312
; Sequence 11312, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11312
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11312

Query Match      88.5% Score 1464; DB 4; Length 341;
Best Local Similarity 87.2%; Pred No. 96-153;
Matches 279; Conservative 16; Mismatches 25; Indels 0; Gaps 0;

QY      1 MPWQVQKHYHGSAFLRVELPTCMYRLPVNHSKTTSPATDAGHVOETSPSLQALESRD 60
        |||||
DB      22 MPWQVQKHYHGSAFLRVELPTCMYRLPVNHSKTTSPATDAGHVOETSPSLQALESRD 81
        |||||

QY      61 DILKLYEKAADVDSKRIHPDADLVNTNLQADEPTTLATNTLDLSVYGYGAK 120
        |||||
DB      82 DILKLYEKAADVDSKRIHPDADLVNTNLQADEPTTLATNTLDLSVYGYGAK 141
        |||||

QY      121 DIVIANPASPPLSLVLRLLCEERYVLTSTHTSSVGVNPEMLVKCFGEQARKOSRAH 180
        |||||
DB      142 DIVIANPASPPLSLVLRLLCEERYVLTSTHTSSVGVNPEMLVKCFGEQARKOSRAH 201
        |||||

QY      181 YOLGFTLLWKVNPQTKQKFSVQTMCPBEGNIAFLFSLFQKNAVATLLIDSVDIA 240
        |||||
DB      202 YOLGFTLLWKVNPQTKQKFSVQTMCPBEGNIAFLFSLFQKNAVATLLIDSVDIA 261
        |||||

QY      241 MQLGEGSKKKAANFRSNLSALGRSPMLVGNELTVADVVLMSVYQQTGSSGAAPTNY 300
        |||||
DB      262 IFQLGEGSKKKAANFRSNLSALGRSPMLVGNELTVADVVLMSVYQQTGSSGAAPTNY 321
        |||||

QY      301 RWLKSCEMLAPSTALQLLK 320
```

Db 322 RWMRSCENTLAPFNATLKLK 341

RESULT 2

US-08-518-862C-2
Sequence 2, Application US/08518862C
Patent No. 5843757
GENERAL INFORMATION:
APPLICANT: Vogelstein, Bert
APPLICANT: Kinzler, Kenneth W.
APPLICANT: Nicolaides, Nicholas C.
TITLE OF INVENTION: Human JTV1 Gene Overlaps PMS2 Gene
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Banner & Witcoff, Ltd.
STREET: 1001 G Street, N.W.
CITY: Washington, D.C.
COUNTRY: U.S.A.
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/518,862C
FILING DATE: 24-AUG-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107,49697
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 312 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-518-862C-2

Query Match 86.8%; Score 1436; DB 2; Length 312;
Best Local Similarity 87.5%; Pred. No. 9,6e-150;
Matches 273; Conservative 14; Mismatches 25; Indels 0; Gaps 0;

QY 1 MEMYQVKPYHGSAPRLVRLPTCMRLPNVHAKTTSPTADAGHVQETSEPSIQALESRQD 60
DB 1 MEMYQVKPYHGSAPRLVRLPTCMRLPNVHGRSYGAPAGAGHVQESNLSIQALESRQD 60
QY 61 DILKRLYELKAANDGSKMHTPPDADLVNTILODEPTLATNTLIDNSVIGKXYGALK 120
DB 61 DILKRLYELKAANDGSKMHTPPDADLVNTILODEPTLATNTLIDNSVIGKXYGALK 120
QY 121 DIVIANPASPPLSLIVHRLICERYRLSTVHTSSVKNVPELVKCFGEQARKQSRRH 180
DB 121 DIVIANPASPPLSLIVHRLICERYRLSTVHTSSVKNVPELVKCFGEQARKQSRRH 180
QY 181 VOLGFTLWKNVPTOMKPSVOTMCPBIGEGNIAPFLSIRGQKKNATVTLIDSWVIA 240
DB 181 VOLGFTLWKNVPTOMKPSVOTMCPBIGEGNIAPFLSIRGQKKNATVTLIDSWVIA 240
QY 241 MFOLREGSSKEKAIVFRSNNALGRSPVLVGNELTVADVILMSVYLQOTGSSGAAPTVO 300
DB 241 IFOLREGSSKEKAIVFRSNNALGRSPVLVGNELTVADVILMSVYLQOTGSSGAAPTVO 300
QY 301 RMLKSCENTLAPF 312
DB 301 RWMRSCENTLAPF 312

RESULT 3

US-09-443-184-48
Sequence 48, Application US/09443184A
Patent No. 6372431
GENERAL INFORMATION:
APPLICANT: Cunningham, Mary Jane
APPLICANT: Zweiger, Gary
APPLICANT: Kaser, Matthew R.
APPLICANT: Panzer, Scott
APPLICANT: Selhammer, Jeffrey J.
APPLICANT: Yue, Henry
APPLICANT: Baughn, Mariah
APPLICANT: Azimzai, Yalda
APPLICANT: Lal, Preeti
TITLE OF INVENTION: MAMMALIAN TOXICOLOGICAL RESPONSE MARKERS
FILE REFERENCE: PC-0007 US
CURRENT APPLICATION NUMBER: US/09/443,184A
CURRENT FILING DATE: 1999-11-19
NUMBER OF SEQ ID NOS: 138
SOFTWARE: PERL Program
SEQ ID NO 48
LENGTH: 1512
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6372431 2302721CD1
US-09-443-184-48

Query Match 6.4%; Score 106; DB 3; Length 1512;
Best Local Similarity 21.8%; Pred. No. 0.096;
Matches 44; Conservative 33; Mismatches 69; Indels 56; Gaps 8;

QY 124 INANPASPPLSLIVHRLICERYRLSTVHTSSVKNVPELVKCFGEQARKQSRRH 183
DB 6 LTVNSGDPPLGALV-----AVEHVKDD-VGISVEGKENTLH----- 41
QY 184 GFTLWKNVPTOMKPSVOTMCPBIGEGNIAPFL-----FSLFGQKKNATVTLIDSWV 237
DB 42 -----VSENVJFTDV-----NSILRYIARVATTAGLYG--NLMETTELIDHWL 82
QY 238 DIAMFOLREGSSKEKAIVFRSNNALGRSPVLVGNELTVADVILMSVYLQ-----QTGSS 292
DB 83 EFSATKL--SSCSFSTITIELNHCLSLRITVIVGNSLSLADLCVWATLKGNAAWQBLKQ 140
QY 293 GAAPTVO RMLKSCENTLAPFST 314
DB 141 KKAIVHVKRWPGFLAQAFQS 162

RESULT 4

US-09-357-251-37
Sequence 37, Application US/09357251
Patent No. 6271441
GENERAL INFORMATION:
APPLICANT: Falco, S. Carl
APPLICANT: Farnodu, Lavo O.
APPLICANT: Orozco, Buddy
APPLICANT: Schwaber, James S.
TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetase
FILE REFERENCE: BB-1193
CURRENT APPLICATION NUMBER: US/09/357,251
CURRENT FILING DATE: 1999-07-20
EARLIER APPLICATION NUMBER: 60/093,530
NUMBER OF SEQ ID NOS: 37
SOFTWARE: Microsoft Office 97
SEQ ID NO 37
LENGTH: 1440
TYPE: PRT
ORGANISM: Homo sapiens
US-09-357-251-37

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:57:26 ; Search time 50.2443 Seconds
(without alignments)
2084.158 Million cell updates/sec

Title: US-10-622-817-6

Perfect score: 1655

Sequence: 1 MPMYQVKRHYGGSAPLRYEL.....RWLKSCEMLAPETALQLLK 320

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*

10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*

11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*

13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*

14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*

15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*

17: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*

18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*

19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|-------------------|
| 1 | 1436 | 86.8 | 312 | US-10-463-676-4 | Sequence 4, Appl1 |
| 2 | 1069 | 64.6 | 229 | US-10-463-676-6 | Sequence 6, Appl1 |
| 3 | 735 | 44.4 | 161 | US-10-463-676-5 | Sequence 5, Appl1 |
| 4 | 105.5 | 6.4 | 925 | US-10-437-963-126132 | Sequence 126132, |
| 5 | 104 | 6.3 | 980 | US-10-369-493-6336 | Sequence 6336, Ap |
| 6 | 103.5 | 6.3 | 948 | US-10-424-599-260388 | Sequence 260388, |
| 7 | 99.5 | 6.0 | 914 | US-10-437-963-153870 | Sequence 153870, |
| 8 | 99 | 6.0 | 691 | US-10-283-132A-44369 | Sequence 44369, A |
| 9 | 98 | 5.9 | 690 | US-09-815-242-5841 | Sequence 5841, Ap |
| 10 | 96.5 | 5.8 | 1440 | US-09-864-464-37 | Sequence 37, Appl |
| 11 | 96.5 | 5.8 | 1440 | US-10-786-720-31 | Sequence 31, Appl |
| 12 | 96 | 5.8 | 569 | US-10-437-963-123904 | Sequence 123904, |
| 13 | 95 | 5.7 | 722 | US-10-369-493-10013 | Sequence 10013, A |

| | | | | | | |
|----|------|-----|------|----|----------------------|-------------------|
| 14 | 93.5 | 5.6 | 456 | 15 | US-10-425-114-63166 | Sequence 63166, A |
| 15 | 93 | 5.6 | 204 | 15 | US-10-369-493-10805 | Sequence 10805, A |
| 16 | 93 | 5.6 | 1398 | 16 | US-10-437-963-180124 | Sequence 180124, |
| 17 | 91 | 5.5 | 201 | 15 | US-10-369-493-13717 | Sequence 13717, A |
| 18 | 91 | 5.5 | 455 | 15 | US-10-425-114-42633 | Sequence 42633, A |
| 19 | 91 | 5.5 | 805 | 15 | US-10-424-599-191796 | Sequence 191796, |
| 20 | 91 | 5.5 | 1319 | 16 | US-10-437-963-197783 | Sequence 197783, |
| 21 | 91 | 5.5 | 1788 | 16 | US-10-437-963-197780 | Sequence 197780, |
| 22 | 90.5 | 5.5 | 419 | 15 | US-10-424-599-260387 | Sequence 260387, |
| 23 | 90.5 | 5.5 | 578 | 15 | US-10-418-8618-32 | Sequence 32, Appl |
| 24 | 90.5 | 5.5 | 661 | 15 | US-10-369-493-2558 | Sequence 2558, Ap |
| 25 | 90.5 | 5.5 | 716 | 15 | US-10-369-493-2175 | Sequence 2175, Ap |
| 26 | 90 | 5.4 | 559 | 9 | US-09-364-847-21 | Sequence 21, Appl |
| 27 | 90 | 5.4 | 856 | 9 | US-09-364-847-33 | Sequence 33, Appl |
| 28 | 90 | 5.4 | 856 | 9 | US-09-364-847-35 | Sequence 35, Appl |
| 29 | 90 | 5.4 | 1500 | 16 | US-10-437-963-111265 | Sequence 111265, |
| 30 | 90 | 5.4 | 5215 | 9 | US-09-861-289-2 | Sequence 2, Appl1 |
| 31 | 90 | 5.4 | 5215 | 9 | US-09-860-846-2 | Sequence 2, Appl1 |
| 32 | 90 | 5.4 | 5215 | 10 | US-09-988-3848-2 | Sequence 2, Appl1 |
| 33 | 90 | 5.4 | 5215 | 10 | US-09-836-821-2 | Sequence 2, Appl1 |
| 34 | 90 | 5.4 | 5215 | 14 | US-10-271-889-45 | Sequence 45, Appl |
| 35 | 89.5 | 5.4 | 203 | 15 | US-10-369-493-8604 | Sequence 8604, Ap |
| 36 | 89 | 5.4 | 407 | 15 | US-10-424-599-267113 | Sequence 267113, |
| 37 | 89 | 5.4 | 1056 | 15 | US-10-282-174-472 | Sequence 472, App |
| 38 | 89 | 5.4 | 1056 | 15 | US-10-282-174-474 | Sequence 476, App |
| 39 | 89 | 5.4 | 1056 | 15 | US-10-282-174-476 | Sequence 476, App |
| 40 | 89 | 5.4 | 1056 | 17 | US-10-600-009-472 | Sequence 472, App |
| 41 | 89 | 5.4 | 1056 | 17 | US-10-600-009-474 | Sequence 474, App |
| 42 | 89 | 5.4 | 1056 | 17 | US-10-600-009-476 | Sequence 476, App |
| 43 | 89 | 5.4 | 1057 | 16 | US-10-714-796-4 | Sequence 4, Appl |
| 44 | 88.5 | 5.3 | 656 | 15 | US-10-369-493-18559 | Sequence 18559, A |
| 45 | 88.5 | 5.3 | 799 | 16 | US-10-437-963-112603 | Sequence 112603, |

ALIGNMENTS

RESULT 1

US-10-463-676-4

Sequence 4, Application US/10463676

Publication No. US20040175375A1

GENERAL INFORMATION:

APPLICANT: Kim, Sunghoon

APPLICANT: Park, Bum-joon

TITLE OF INVENTION: Method for Treating Cancer Using P38/JTV-1 and Method

FILE REFERENCE: 012679-091

CURRENT APPLICATION NUMBER: US/10/463,676

CURRENT FILING DATE: 2003-06-18

PRIOR APPLICATION NUMBER: KR 10-2003-13058

PRIOR FILING DATE: 2003-03-03

NUMBER OF SEQ ID NOS: 17

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 4

LENGTH: 312

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: PEPTIDE

LOCATION: (1)...(312)

OTHER INFORMATION: 1-312 amino acid sequence of p38/JTV-1

US-10-463-676-4

Query Match 86.8%; Score 1436; DB 16; Length 312;

Best Local Similarity 87.5%; Pred. No. 2.6e-137;

Matches 273; Conservative 14; Mismatches 25; Indels 0; Gaps 0;

QY 1 MPMYQVKRHYGGSAPLRYELPCTMRPLPVHSGKTSPTADAGHVOETSPSSLOAESROD 60

DB 1 MPMYQVKRHYGGSAPLRYELPCTMRPLPVHSGKTSPTADAGHVOETSPSSLOAESROD 60

QY 61 DILKLEIKAVDGLSKVHITPDADLVNTILQADEPTLATNTLDLSVLGKDYGALK 120

```

Db      6  DLKRLVELKAAVDGSLXMIQTPEDADLVNTNIIOADEPTTLTNAIDLNSVKGDOYALK 120

Qy      121  DIVYNNANPAPPRLSLVYHRLLCERYVYLVTHVHSSVKNVPEMLVKCFGEAROKSRHE 180
        121  DIVYNNANPAPRLSLVYHRLLCERHFRVYLVTHVHSSVKNVPEMLVKCFGEAROKSRH 180
Db      181  YOLGFTLIMWNVNPKTOMKFSVOTMCEBEGGNIAFLSLFGOKHNANTLTLIDSWDIA 240
        181  YOLGFTLIMWNVNPKTOMKFSVOTMCEBEGGNIAFLSLFGOKHNANNTLTLIDSWDIA 240
Qy      181  YOLGFTLIMWNVNPKTOMKFSVOTMCEBEGGNIAFLSLFGOKHNANTLTLIDSWDIA 240
        181  YOLGFTLIMWNVNPKTOMKFSVOTMCEBEGGNIAFLSLFGOKHNANNTLTLIDSWDIA 240
Db      241  MFOLRBGSSGEKKAAYFRSMNSALGKSPMLVGNELTVADVILSVLQOTGSGSGAAPTNYQ 300
        241  MFOLRBGSSGEKKAAYFRSMNSALGKSPMLVGNELTVADVILSVLQOTGSGSGAAPTNYQ 300
Qy      241  IFOLKGGSSGEKKAAYFRSMNSALGKSPMLVGNELTVADVILSVLQOTGSGSGAAPTNYQ 300
        241  IFOLKGGSSGEKKAAYFRSMNSALGKSPMLVGNELTVADVILSVLQOTGSGSGAAPTNYQ 300
Db      301  RMLKSCENLAPF 312
        301  RMLKSCENLAPF 312
Qy      301  RMLKSCENLAPF 312
        301  RMLKSCENLAPF 312

```

```

RESULT 2
US-10-463-676-6
; Sequence 6, Application US/10463676
; Publication No. US20040175375A1
GENERAL INFORMATION:
APPLICANT: Kim, Sunghoon
APPLICANT: Park, Bum-Joon
TITLE OF INVENTION: Method for Treating Cancer using p38/JTV-1 and Method
of Invention: for Screening Pharmaceutical Composition for Treating Cancer
FILE REFERENCE: 012679-091
CURRENT APPLICATION NUMBER: US/10/463,676
PRIORITY FILING DATE: 2003-06-18
PRIORITY APPLICATION NUMBER: KR 10-2003-13058
PRIORITY FILING DATE: 2003-03-03
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 229
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (1)...(229)
OTHER INFORMATION: 84-312 amino acid sequence p38/JTV-1
US-10-463-676-6

```

| | | | | |
|-----------------------|-------|--|---------------|------------|
| Query Match | 64.6% | Score 1069 | DR 16 | Length 229 |
| Best Local Similarity | 88.2% | Pred. No. 46-100 | | |
| Matches | 202 | Conservative 12 | Mismatches 15 | Indels 0 |
| | | | Gaps | 0 |
| QY | 84 | DADLDVTNILOADEPTLTATNTLIDNSVLGDKYGLKDIVINANASPELSLTVLRLLC | 143 | |
| Db | 1 | DADLDVTNILOADEPTLTATNTLIDNSVLGDKYGLKDIVINANASPELSLTVLRLLC | 60 | |
| QY | 144 | ERYRVLSTVTHSSSVKNVPENLYCFCFGQAKRKQSHHEVQLGFTLLIKWNPPTOMKEFSVOT | 203 | |
| Db | 61 | EHFRVLSTVTHSSSVKSVPEMLKKCFGQONKKOPRODVLPTLLIKWNPPTOMKEFSIOT | 120 | |
| QY | 204 | MCPIEGEGNIAFLFSLFGQKHNATVTLTIDSWDIAFMFQUREGSSKEKAAPFSNSAL | 263 | |
| Db | 121 | MCPLEEGEGNIAFLFSLFGQKHNATVTLTIDSWDIAFMFQUREGSSKEKAAPFSNSAL | 180 | |
| QY | 264 | GRSPMLVGNELTIVADVYLMSTVLOQTGGSSGAAPITNVQWLKSCENLAPF | 312 | |
| Db | 181 | GKSPMLVGNELTIVADVYLMSTVLOQTGGSSVTPANVRWMSKCNLAPF | 229 | |

RESULT3
US-10-463-676-5
; Sequence 5, Application US/10463676
; Publication No. US20040175375A1
; GENERAL INFORMATION:
; APPLICANT: Klm, Sungmoon
; APPLICANT: Park, Bnn-Joon

```

:
: TITLE OF INVENTION: Method for Treating Cancer Using p38/JTV-1 and Method
: TITLE OF INVENTION: for Screening Pharmaceutical Composition for Treating Cancer
: FILE REFERENCE: 01679-091
: CURRENT APPLICATION NUMBER: US/10/463, 676
: CURRENT FILING DATE: 2003-06-18
: PRIOR APPLICATION NUMBER: KR 10-2003-13058
: PRIOR FILING DATE: 2003-03-03
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 5
: LENGTH: 161
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: PEPTIDE
: LOCATION: (1)...(161)
: OTHER INFORMATION: 1-161 amino acid sequence of p38/JTV-1
: US-10-463-676-5

```

| Query Match | 44.4% | Score 735 | DB 16 | Length 161 |
|-----------------------|----------------|---|----------|------------|
| Best Local Similarity | 88.8% | Pred. No. 2.5e-66 | | |
| Matches 143 | Conservative 5 | Mismatches 13 | Indels 0 | Gaps 0 |
| QY | 1 | MPMYQVKPYHGGGAPLRLVELPTCYMYRLPNVHSKTTSEATDAGVQETSEPSLOALESROD | 60 | |
| Db | 1 | MPMYQVKPYHGGGAPLRLVELPTCYMYRLPNVHGSYGPAPGAVQESNLSIQALESROD | 60 | |
| QY | 61 | DLTKRLRYELKAAVDSLSMITHTPDADLDVNNIIQADEPTTLATNTILNLSVIGKDYGLK | 120 | |
| Db | 61 | DLTKRLRYELKAAVDSLSMITHTPDADLDVNNIIQADEPTTLATNTILNLSVIGKDYGLK | 120 | |
| QY | 121 | DIVINANPASPPLSLVLVHRLCGRYVLSVTYHSSVKNV | 161 | |
| Db | 121 | DIVINANPASPPLSLVLVHRLCGRFVLSVTYHSSVKS | 161 | |

RESULT 4
 US-10-437-963-126132
 ; Sequence 126132, Application US/10437963
 ; Publication No. US2004012345A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21 (53221) B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; NUMBER OF SEQ ID NOS: 2003-05-14
 ; SEQ ID NO 126132
 ; LENGTH: 925
 ; TYPE: PRT
 ; ORGANISM: *Oryza sativa*
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_28709C.1.dep
 US-10-437-963-126132

| | | | | |
|-----------------------|-------|--|----------------|------------|
| Query Match | 6.4% | Score 105.5 | DB 16 | Length 925 |
| Best Local Similarity | 20.3% | Pred. No. 0.51 | | |
| Matches | 64 | Conservative 46 | Mismatches 107 | Indels 99 |
| | | | | Gaps 13 |
| QY | 25 | YELPNVHSKTTSPATDAGHVOETSEPTQLAESRODDI-LKRLVELKAAVGLSKMHT | 82 | |
| | | | | |
| Db | 152 | YVVDNIIRAKTGPF-----TIDPRLQLLYKKTTELVGIEKQSKVLKILSLSDVHA | 203 | |
| | | | | |
| QY | 83 | PDADLVNIIQAD-EPTLTATITDLSNVLGKDYGA-----LKDVIIVAN | 127 | |
| | | | | |

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model.

Run on: February 23, 2005, 13:34:50 / Search time 13.3213 Seconds
(without alignments)
2311.294 Million cell updates/sec

Title: US-10-622-817-6

Perfect score: 1655

Sequence: 1 MPMTQVKKPHGSGAPRLVRL.....RWLKSCEMLAPSTALQLLK 320

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: pir79:*
2: pir12:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 126.5 | 7.6 | 719 | 2 | T52043 |
| 2 | 112 | 6.8 | 728 | 2 | T01200 |
| 3 | 104 | 6.3 | 980 | 2 | T30089 |
| 4 | 99 | 6.0 | 691 | 2 | B89797 |
| 5 | 98.5 | 6.0 | 222 | 2 | B82441 |
| 6 | 98.5 | 6.0 | 377 | 2 | F69172 |
| 7 | 98 | 5.9 | 690 | 2 | A24545 |
| 8 | 96.5 | 5.8 | 1440 | 1 | STRUQT |
| 9 | 95.5 | 5.8 | 753 | 2 | T24869 |
| 10 | 95 | 5.7 | 488 | 1 | H64313 |
| 11 | 95 | 5.7 | 552 | 2 | D82421 |
| 12 | 95 | 5.7 | 2655 | 2 | D96595 |
| 13 | 94.5 | 5.7 | 2471 | 2 | T42977 |
| 14 | 94 | 5.7 | 435 | 2 | T01567 |
| 15 | 94 | 5.7 | 436 | 2 | C86719 |
| 16 | 93.5 | 5.6 | 1064 | 2 | B86465 |
| 17 | 93 | 5.6 | 490 | 2 | B84999 |
| 18 | 92 | 5.6 | 754 | 2 | S37403 |
| 19 | 91.5 | 5.5 | 436 | 2 | T51237 |
| 20 | 91 | 5.5 | 1265 | 2 | T47626 |
| 21 | 90.5 | 5.5 | 359 | 2 | T37921 |
| 22 | 90.5 | 5.5 | 456 | 2 | C84764 |
| 23 | 90.5 | 5.5 | 642 | 2 | T39490 |
| 24 | 90.5 | 5.5 | 716 | 2 | T37830 |
| 25 | 90 | 5.4 | 437 | 2 | T51238 |
| 26 | 90 | 5.4 | 559 | 2 | A38604 |
| 27 | 89.5 | 5.4 | 2954 | 2 | T14156 |
| 28 | 89.5 | 5.4 | 317 | 2 | T05528 |
| 29 | 89.5 | 5.4 | 1597 | 2 | S65053 |

| | | | | | | |
|----|------|-----|------|---|--------|--------------------|
| 30 | 89.5 | 5.4 | 1714 | 1 | S18644 | multifunctional am |
| 31 | 89.5 | 5.4 | 4976 | 2 | T14165 | peptide synthetase |
| 32 | 89 | 5.4 | 184 | 2 | S74356 | glutathione S-tran |
| 33 | 89 | 5.4 | 763 | 2 | A62443 | penicillin-binding |
| 34 | 89 | 5.4 | 1023 | 2 | T31669 | neural zinc finger |
| 35 | 89 | 5.4 | 1056 | 1 | G02157 | kinesin-like spind |
| 36 | 89 | 5.4 | 1247 | 2 | A33812 | interphotoreceptor |
| 37 | 88.5 | 5.3 | 602 | 2 | A83542 | gtp-binding protei |
| 38 | 88.5 | 5.3 | 656 | 2 | H84206 | acetyl-CoA synthet |
| 39 | 88 | 5.3 | 821 | 2 | T24728 | hypothetical prote |
| 40 | 88 | 5.3 | 2279 | 2 | T42531 | acetyl-CoA carboxy |
| 41 | 87.5 | 5.3 | 2280 | 2 | T38906 | acetyl-CoA carboxy |
| 42 | 87.5 | 5.3 | 474 | 2 | B70597 | hypothetical prote |
| 43 | 87.5 | 5.3 | 820 | 2 | H82302 | ATP-dependent heli |
| 44 | 87 | 5.3 | 570 | 2 | T15763 | hypothetical prote |
| 45 | 86 | 5.2 | 906 | 2 | T45158 | pre-mRNA splicing |

ALIGNMENTS

RESULT 1

T52043

probable glutamate-tRNA ligase (EC 6.1.1.17) [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004

C:Accession: T52043

R:Day, T.S.; Golovkin, M.; Reddy, A.S.

Biochim. Biophys. Acta 1399, 219-24, 1998

A>Title: Cloning of the cDNA for glutamyl-tRNA synthetase from Arabidopsis thaliana.

A:Reference number: 224836; PMID:9765600; PMID:9765600

A:Accession: T52043

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-719 <DAY>

A:Cross-references: UNIPROT:O82462; EMBL:AF067773; PDB:1AC36469.1

C:Superfamily: Yeast glutamate-tRNA ligase; glutamine-tRNA ligase homology

C:Keywords: ligase

Query Match 7.6%; Score 126.5; DB 2; Length 719;

Best Local Similarity 28.0%; Pred. No. 0.017;

Matches 49; Conservative 22; Mismatches 57; Indels 47; Gaps 8;

QY 128 PASPLSLVLRLLCEYRVLTSTHTSSVKNVPELVKCRGEQAKRSRREYQLGFL 187

DB 10 PESPLSVLVALSLASPV---TTDSSAAATTVPSFV---SPGRKLN-----GATV 55

QY 188 IWKVPTQMKFSVQTMCPTEEGENIARFLFSLFGOKHNAVTLTLDVWDIAMFQLREG 247

DB 56 LLRYV-----GRSAKKLPDFVG--NNAPDSQIDENVDAVSV-FSSG 94

QY 248 SSKRAVPRSMNSALGRSPMLVGNELTVADVLMVLTQCTGSSGAAPTQVRW 302

DB 95 SEFNAC--GRVDKLTSLFVGHSLIADVAIWSALAGTG-----QRW 137

RESULT 2

T01200

probable glutamate-tRNA ligase (EC 6.1.1.17) F21E10.12 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004

C:Accession: T01200

R:Davidson, S.; Rohlfing, T.; David, M.; O'Brian, D.

Submitted to the EMBL Data Library, April 1998

A>Description: The sequence of A. thaliana F21E10.

A:Reference number: 214258

A:Accession: T01200

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-728 <DAY>

A:Cross-references: UNIPROT:O65253; EMBL:AF058914; NID:G3047074; PID:G3047084; GSPDB:GNO

A:Experimental source: cultivar Columbia

C:Genetics:

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:33:04 ; Search time 60.8145 Seconds
(without alignments)
2694.512 Million cell updates/sec

Title: US-10-622-817-6
Perfect score: 1655
Sequence: 1 MPWQVPRHYGSAFLRVRL.....RWLXSCENLAPFTALQLLK 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 03:*
1: uniprot_sprot:*
2: uniprot_crembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|--------------|---------------------|
| 1 | 1655 | 100.0 | 320 | 08R010 | 08R010 mus musculu |
| 2 | 1652 | 99.8 | 320 | 08R2Y6 | 08R2Y6 mus musculu |
| 3 | 1573 | 95.0 | 320 | 1 MCA2_CRIGR | 09wrm7 cricetulus |
| 4 | 1464 | 88.5 | 320 | 1 MCA2_HUMAN | 013155 homo sapien |
| 5 | 1460 | 88.2 | 320 | 2 096CZ5 | 096CZ5 homo sapien |
| 6 | 1412 | 85.3 | 280 | 2 08R3V2 | 08R3V2 mus musculu |
| 7 | 1153.5 | 69.7 | 311 | 2 06DK86 | 06DK86 xenopus tro |
| 8 | 1146.5 | 69.3 | 311 | 2 06TNU4 | 06TNU4 xenopus lae |
| 9 | 1146.5 | 69.3 | 311 | 2 07ZXD7 | 07ZXD7 xenopus lae |
| 10 | 926.5 | 56.0 | 321 | 2 0773C0 | 0773C0 brachydantio |
| 11 | 247 | 14.9 | 340 | 2 07Q7A3 | 07Q7A3 anopheles g |
| 12 | 224 | 13.5 | 301 | 2 07KUM5 | 07KUM5 drosophila |
| 13 | 224 | 13.5 | 313 | 2 08T060 | 08T060 drosophila |
| 14 | 224 | 13.5 | 322 | 2 06NKM4 | 06NKM4 drosophila |
| 15 | 216.5 | 13.1 | 334 | 1 MCA2_DROME | 09vur3 drosophila |
| 16 | 126.5 | 7.6 | 719 | 2 082462 | 082462 arabidopsis |
| 17 | 112 | 6.8 | 728 | 2 065253 | 065253 arabidopsis |
| 18 | 107 | 6.5 | 715 | 2 09LJ28 | 09LJ28 oryza sativ |
| 19 | 106 | 6.4 | 328 | 2 086X73 | 086X73 homo sapien |
| 20 | 106 | 6.4 | 559 | 2 08RPZ6 | 08RPZ6 pseudomons |
| 21 | 106 | 6.4 | 869 | 2 06PD57 | 06PD57 homo sapien |
| 22 | 106 | 6.4 | 1721 | 2 07NID9 | 07NID9 gloeobacter |
| 23 | 104 | 6.3 | 980 | 2 017592 | 017592 caenorhabdi |
| 24 | 102.5 | 6.2 | 420 | 2 08RFOO | 08RFOO corynebacte |
| 25 | 102 | 6.2 | 559 | 2 0939A8 | 0939A8 pseudomons |
| 26 | 102 | 6.2 | 1048 | 2 07N3E2 | 07N3E2 photoxabbu |
| 27 | 101.5 | 6.1 | 715 | 2 091B77 | 091B77 white spot |
| 28 | 101.5 | 6.1 | 730 | 2 0910F5 | 0910F5 white spot |
| 29 | 101.5 | 6.1 | 913 | 2 06DRB3 | 06DRB3 brachydantio |
| 30 | 101.5 | 6.1 | 913 | 2 06PFO4 | 06PFO4 brachydantio |
| 31 | 100.5 | 6.1 | 555 | 2 07MEV7 | 07MEV7 vibrio vuln |

| | | | | | | |
|----|-------|-----|------|---|-----------|---------------------|
| 32 | 100.5 | 6.1 | 555 | 2 | 08D7U3 | 08D7U3 vibrio vuln |
| 33 | 100.5 | 6.1 | 922 | 2 | 08K2J4 | 08K2J4 mus musculu |
| 34 | 99.5 | 6.0 | 885 | 2 | 06REMS | 06REMS brachydantio |
| 35 | 99.5 | 6.0 | 559 | 2 | 08R067 | 08R067 pseudomons |
| 36 | 99 | 6.0 | 690 | 1 | LIP_STRAW | 08NYC2 staphylococ |
| 37 | 99 | 6.0 | 690 | 2 | 06GCF1 | 06GCF1 staphylococ |
| 38 | 99 | 6.0 | 691 | 2 | 0798Z7 | 0798Z7 staphylococ |
| 39 | 99 | 6.0 | 691 | 2 | 09W06 | 09W06 staphylococ |
| 40 | 99 | 6.0 | 691 | 2 | 07A7P2 | 07A7P2 staphylococ |
| 41 | 98.5 | 6.0 | 222 | 2 | 09KM05 | 09KM05 vibrio chol |
| 42 | 98.5 | 6.0 | 377 | 2 | 026161 | 026161 methanobact |
| 43 | 98.5 | 6.0 | 1597 | 2 | 08V9A2 | 08V9A2 crucifer to |
| 44 | 98 | 5.9 | 240 | 2 | 08BTR1 | 08BTR1 mus musculu |
| 45 | 98 | 5.9 | 690 | 1 | LIP_STRAW | P10335 staphylococ |

ALIGNMENTS

```
RESULT 1
08R010
ID 08R010; PRELIMINARY; PRT: 320 AA.
AC 08R010;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Utrv-pending protein.
GN Name=Utrv;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N, and FVB/N-3; TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnae.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Altschul S.F., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Matulska K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Mullahy S.J.,
RA Raha S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Boeak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny K.C., Harte S., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
RA Jones S.J., Mair M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA Strausberg R.;
RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N-3; TISSUE=Mammary tumor;
RA Strausberg R.;
RT Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024410; AAH24410.1; -.
DR EMBL; BC026972; AAH26972.1; -.
DR MGI; MGI:2385237; Utrv1.
DR InterPro; IPR004046; GST_C-term.
DR InterPro; IPR010987; GST_C-like.
DR Pfam; PF00043; GST_C; 1.
SQ SEQUENCE 320 AA; 35396 MW; 1745D7BE4BC3670D CRC64;
```

| Query Match | Similarity | 100.0% | Score | 1655 | DB 2 | Length | 320 |
|-------------|--|--|-----------|------------|------|--------|------|
| Best Local | Similarity | 100.0% | Pred. No. | 2,6e-128 | | | |
| Matches | 320 | Conservative | 0 | Mismatches | 0 | Indels | 0 |
| | | | | | | | Gaps |
| | | | | | | | 0 |
| QY | 1 | MEMYQVKKPHGGGAPRLVELPTCMYTLPNVHSKTTSPATAGAGVOETSEPSLOALSSROD | 60 | | | | |
| Db | 1 | MEMYQVKKPHGGGAPRLVELPTCMYTLPNVHSKTTSPATAGAGVOETSEPSLOALSSROD | 60 | | | | |
| QY | 61 | DILKRLYEYKAADVGSKNIHTPPDADLDVTNIIQADEPTTLATNTLIDNSVIGKDYGALK | 120 | | | | |
| Db | 61 | DILKRLYEYKAADVGSKNIHTPPDADLDVTNIIQADEPTTLATNTLIDNSVIGKDYGALK | 120 | | | | |
| QY | 121 | DIVINANPASPPLSLIVLHRLCERRVLYSTVHTHSYKRVNPENLYKCGEORAKGRHE | 180 | | | | |
| Db | 121 | DIVINANPASPPLSLIVLHRLCERRVLYSTVHTHSYKRVNPENLYKCGEORAKGRHE | 180 | | | | |
| QY | 181 | YOLGFLTLMKNVPKYOMKPSVOTMCPIDEBGNIAFLFSLFGQKNAVTLTLIDSVWDIA | 240 | | | | |
| Db | 181 | YOLGFLTLMKNVPKYOMKPSVOTMCPIDEBGNIAFLFSLFGQKNAVTLTLIDSVWDIA | 240 | | | | |
| QY | 241 | MFQLREGSSKEKRAVFRSNNSALGRPMTLVGNELTADVAVLMSVLOQTGSSGGAAPTNO | 300 | | | | |
| Db | 241 | MFQLREGSSKEKRAVFRSNNSALGRPMTLVGNELTADVAVLMSVLOQTGSSGGAAPTNO | 300 | | | | |
| QY | 301 | FWLKSCEMLAPSTALQOLIK | 320 | | | | |
| Db | 301 | FWLKSCEMLAPSTALQOLIK | 320 | | | | |
| RESULT 2 | | | | | | | |
| QY | 08R2Y6 | PRELIMINARY | PRT | 320 | AA | | |
| AC | 08R2Y6 | | | | | | |
| DT | 01-JUN-2002 | (TREMBlrel. 21, Created) | | | | | |
| DT | 01-JUN-2002 | (TREMBlrel. 21, Last sequence update) | | | | | |
| DT | 01-MAR-2004 | (TREMBlrel. 26, Last annotation update) | | | | | |
| DE | YJY1 | pending protein. | | | | | |
| GN | Name=YJY1 | | | | | | |
| OS | Mus musculus (Mouse). | | | | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | | | |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | | | | |
| NCBI | Taxid=10090; | | | | | | |
| | [1] | | | | | | |
| RP | SEQUENCE FROM N.A. | | | | | | |
| RP | STRAIN=CZECH II; TISSUE=Mammary tumor; | | | | | | |
| RC | MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899; | | | | | | |
| RC | Strausberg R.L., Feingold E.A., Grove L.H., Derge J.G., | | | | | | |
| RA | Klausner R.D., Collins F.S., Wagner L., Shemien C.M., Schuler G.D., | | | | | | |
| RA | Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., | | | | | | |
| RA | Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., | | | | | | |
| RA | Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., | | | | | | |
| RA | Stadlton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., | | | | | | |
| RA | Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C., | | | | | | |
| RA | Rata S.S., Loguella N.A., Peters G.J., Abramson R.D., Mullaly S.J., | | | | | | |
| RA | Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H., | | | | | | |
| RA | Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W., | | | | | | |
| RA | Vallatou D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., | | | | | | |
| RA | Faney J., Helton E., Kettelman M., Madan A.C., Rodriguez S., Sanchez A., | | | | | | |
| RA | Whitney R.M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., | | | | | | |
| RA | Blakeley R.C., Grimwood J., Schmutz J., Myers R.M., Butcherfield Y.S., | | | | | | |
| RA | Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E., | | | | | | |
| RA | Jones S.J., Marra M.A.; | | | | | | |
| RT | "Generation and initial analysis of more than 15,000 full-length human | | | | | | |
| RT | and mouse cDNA sequences"; | | | | | | |
| RL | Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903(2002). | | | | | | |
| RL | [2] | | | | | | |
| RP | SEQUENCE FROM N.A. | | | | | | |
| RC | STRAIN=CZECH II; TISSUE=Mammary tumor; | | | | | | |
| RC | Strausberg R.; | | | | | | |
| RA | Submitted (Apr.-2002) to the EMBL/GenBank/DBJ databases. | | | | | | |
| RA | EMBL; BC026958; AA326958.1; -. | | | | | | |

```

DR      MGD; MG1:2385237; Jctv1.
DR      InterPro; IPR004046; GST_Cterm.
DR      InterPro; IPR010987; GST_C_1like.
DR      Pfam; PF00043; GST_C; 1.
DR      SEQUENCE   320 AA;  35423 MW;  1C21F1A74C9882B4 CRC64;

Query Match      99.8%; Score 1652; DB 2; Length 320;
Best Local Similarity 99.7%; Pred. No. 4,6e-128;
Matches 319; Conservative 1; Mismatches 0; Indels 0; Gaps 0

QY      1 MEMVQVYPHGSGAPLRLVETPTCMYRLPMVHKKTTSPADAGHVQETSEPSIQALESRD 60
DB      1 MEMVQVYPHGSGAPLRLVETPTCMYRLPMVHKKTTNPADAGHVQETSEPSIQALESRD 60
QY      61 DILKRLYEIKAAVDGSKRIHTPPDADLVNTILQADEPTTLATNTLDLSVYKGYGALK 120
DB      61 DILKRLYEIKAAVDGSKRIHTPPDADLVNTILQADEPTTLATNTLDLSVYKGYGALK 120
QY      121 DIVINANPASPPLSLVLVHRLLCERYRVLTSTVHTSSVKNVPENLYKCGEQARKQSRHE 180
DB      121 DIVINANPASPPLSLVLVHRLLCERYRVLTSTVHTSSVKNVPENLYKCGEQARKQSRHE 180
QY      181 YOLGFTLLWKQVPTQMKPSVOTWCPRIEGEINIAFLPSLFGQKNAVTLTLDISWVDA 240
DB      181 YOLGFTLLWKQVPTQMKPSVOTWCPRIEGEINIAFLPSLFGQKNAVTLTLDISWVDA 240
QY      241 MFQLREGSKKEAAYFRSMNALSGLSPMLVNGELTVAADVLSVYLTQGTGGSGAAPPTNVQ 300
DB      241 MFQLREGSKKEAAYFRSMNALSGLSPMLVNGELTVAADVLSVLTQGTGGSGAAPPTNVQ 300
QY      301 RWLXSCENIAPFPSTALQLLK 320
DB      301 RWLXSCENIAPFPSTALQLLK 320

RESULT 3
KCA2_CRIGR
KCA2_CRIGR STANDARD; PRT; 320 AA.
AC      Q9WMT7;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      25-OCT-2004 (Rel. 45, Last annotation update)
DE      Multisynthetase complex auxiliary component p38.
OS      Cricetulus griseus (Chinse hamster).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC      Cricetulus.
RX      NCBI_TaxID=10029;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Ovary;
RX      MEDLINE=99096915; PubMed=9878398; DOI=10.1006/jmbi.1998.2316;
RA      Quevillon S., Robinson J.-C., Berthomieu B., Siatecka M., Mirande M.;
RT      "Macromolecular assemblage of aminocycl-tRNA synthetases:
RT      identification of protein-protein interactions and characterization of
RT      a core protein."
RL      J. Mol. Biol. 285:183-195(1999).
CC      -1- FUNCTION: Probable core protein of the multisynthetase complex
CC      that serves as a template for the assembly of the supramolecular
CC      structure.
CC      -1- SUBUNIT: Component of the multisynthetase complex which is
CC      comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the
CC      monospecific isoleucyl, leucyl, glutamyl, methionyl, lysyl,
CC      arginyl, and aspartyl-tRNA synthetases as well as three auxiliary
CC      proteins, p18, p48 and p43.
CC      -1- SIMILARITY: Contains 1 GST-like domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.1sb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF072727; AAD34422.1; -.
CC InterPro: IPR010987; GST_C_1like.
CC InterPro: IPR004046; GST_Cterm.
CC Pfam: PF00043; GST_C.1.
CC Protein biosynthesis.
CC SEQUENCE 320 AA; 35433 MW; 6D24E03ABEC610A CRC64;

Query Match 95.0%; Score 1573; DB 1; Length 320;
Best Local Similarity 94.4%; Pred. No. 1.6e-121;
Matches 302; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

Qy 1 MPMYGVKPYHGSGAPLRVELPTCMYRLPNVHSKTTSPATDAGHVOETSPSIOALESROD 60
Db 1 MPMYGVKPYHGSGAPLRVELPTCMYRLPNVHSKTTSPATDAGHVOETSPSIOALESROD 60
Qy 61 DILKRLYEKKAADVGLSKMHTPDADLVNTIIQADEPTTLATNTLDINSVLGKDYGLK 120
Db 61 DILKRLYEKKAADVGLSKMHTPDADLVNTIIQADEPTTLATNTLDINSVLGKDYGLK 120
Qy 121 DIVINANASPPSLILVHRLICERYRLSTVHTSSVKNVBNLVKCGEQARKQSRHE 180
Db 121 DIVINANASPPSLILVHRLICERYRLSTVHTSSVKNVBNLVKCGEQARKQSRHE 180
Qy 181 YOLGFTLIMKNVPTQMFKFSVQTMCPIGEGNIARFSLFGCKNIVTLTLIDSWVIA 240
Db 181 YOLGFTLIMKNVPTQMFKFSVQTMCPIGEGNIARFSLFGCKNIVTLTLIDSWVIA 240
Qy 241 MFOLEGGSSKEKAFAVRSNMNSALGRSPMLVGNELTVADVILMSVLQOTGSSGAAPTVO 300
Db 241 MFOLEGGSSKEKAFAVRSNMNSALGRSPMLVGNELTVADVILMSVLQOTGSSGAAPTVO 300
Qy 301 RWLKSCEMLAPSTMLQLK 320
Db 301 RWLKSCEMLAPSTMLQLK 320
Qy 301 RWLKSCEMLAPSTMLQLK 320
Db 301 RWLKSCEMLAPSTMLQLK 320

RESULT 4
MCA2_HUMAN STANDARD; PRT; 320 AA.
ID MCA2_HUMAN
AC Q13155; Q9P1L2;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Multisynthetase complex auxiliary component p38 (JTV-1 protein)
DE (PRO00992).
GN Name=JTV1;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9615582; PubMed=8666379;
RA Nicolaiides N.C., Kinzler K.W., Vogelstein B.;
RT "Analysis of the 5' region of PMS2 reveals heterogeneous transcripts
RT and a novel overlapping gene.";
RL Genomics 29:329-334(1995).
RP [2]
RP SEQUENCE FROM N.A.
RP TISSUE=Lymph;
RX MEDLINE=2338857; PubMed=12477932; DOI=10.1073/pnae.242603899;
RA Straubeberg R.V., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeebber B., Bueter K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Datchenko L., Marnuska K., Farmer A.A., Rubin J., Hong L.,
RA Stepien M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedon T.B., Toshiyuki S., Carrinck P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Guaratine P.H.,
RA Richard S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hilyk S.W.,

```

```

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.W., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE OF 197-320 FROM N.A.
RP TISSUE=Fetal Liver;
RA Zhang C., Yu Y., Zhang S., Wei H., Zhou G., Ouyang S., Luo L., Bi J.,
RA Liu M., He F.;
RT "Functional prediction of the coding sequences of 121 new genes
RT deduced by analysis of cDNA clones from human fetal liver.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP INTERACTION WITH FUBP1.
RX MEDLINE=22716800; PubMed=12819782; DOI=10.1038/ng1182;
RA Kim M.J., Park B.-J., Kang Y.-S., Kim H.-J., Park J.-H., Kang J.W.,
RA Lee S.W., Han J.M., Lee H.-W., Kim S.;
RT "Downregulation of FUBS-binding protein and c-myc by RNA synthetase
RT cofactor p38 is required for lung cell differentiation.";
RL Nat. Genet. 34:330-336(2003).
CC -1- FUNCTION: Probable core protein of the multisynthetase complex
CC that serves as a template for the assembly of the supramolecular
CC structure. Mediates ubiquitination of FUBP1 and its degradation by
CC the proteasome.
CC -1- SUBUNIT: Component of the multisynthetase complex which is
CC comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the
CC monospecific isoleucyl, leucyl, glutamyl, methionyl, lysyl,
CC arginyl, and aspartyl-tRNA synthetases as well as three auxiliary
CC proteins, p18, p48 and p43. Binds FUBP1.
CC -1- SIMILARITY: Contains 1 GST-like domain.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 312.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U24169; AAC50391.1; ALT FRAME.
DR EMBL: BC002853; AAH02853.1; -.
DR EMBL: BC010156; AAH01056.1; -.
DR EMBL: AF116615; AAF71039.1; -.
DR H-InvDB: HIX006460; -.
DR MIM: 600859; -.
DR InterPro: IPR010987; GST_C_1like.
DR InterPro: IPR004046; GST_Cterm.
DR Pfam: PF00043; GST_C.1.
KM Protein biosynthesis.
SQ SEQUENCE 320 AA; 35349 MW; F253726B63C12BAB CRC64;

Query Match 88.5%; Score 1464; DB 1; Length 320;
Best Local Similarity 87.2%; Pred. No. 1.6e-112;
Matches 279; Conservative 16; Mismatches 25; Indels 0; Gaps 0;

Qy 1 MPMYGVKPYHGSGAPLRVELPTCMYRLPNVHSKTTSPATDAGHVOETSPSIOALESROD 60
Db 1 MPMYGVKPYHGSGAPLRVELPTCMYRLPNVHSKTTSPATDAGHVOETSPSIOALESROD 60
Qy 61 DILKRLYEKKAADVGLSKMHTPDADLVNTIIQADEPTTLATNTLDINSVLGKDYGLK 120
Db 61 DILKRLYEKKAADVGLSKMHTPDADLVNTIIQADEPTTLATNTLDINSVLGKDYGLK 120
Qy 121 DIVINANASPPSLILVHRLICERYRLSTVHTSSVKNVBNLVKCGEQARKQSRHE 180
Db 121 DIVINANASPPSLILVHRLICERYRLSTVHTSSVKNVBNLVKCGEQARKQSRHE 180

```


Db 121 DIVINANPASPPLSLVLHRLLCHEFRVLTSTVHTSSVKSPENLTKCPGQNKQKOPROD 180
 Qy 181 YOLGFTLLWKVVPKTKQMFVSQVTCPIEGEGNINARFLFSLFGQKNAVTLTLIDSWDIA 240
 Db 181 YOLGFTLLWKVVPKTKQMFVSQVTCPIEGEGNINARFLFSLFGQKNAVTLTLIDSWDIA 240
 Qy 241 MFOLREGSSKEKAFAVFSMSNLSALGRSPMLVGNELTVADVLMVSLVQOTGGSSGAAPPTNVQ 300
 Db 241 IFOLKEGSSKEKAFAVFSMSNLSALGRSPMLVGNELTVADVLMVSLVQOTGGSSGAAPPTNVQ 300
 Qy 301 RMLKSCENIAPFPSTALQLLK 320
 Db 301 RMLKSCENIAPFPSTALQLLK 320

RESULT 5
 096CZ5 PRELIMINARY; PRT; 320 AA.

AC 096CZ5; 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE JTV1.
 GN Name=JTV1;
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalek U., Smallos D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA Strausberg R.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC013630; AAH13630.1; -;
 DR InterPro; IPR004046; GST_Cterm.
 DR InterPro; IPR010987; GST_C-like.
 DR Pfam; PF00043; GST_C_1_-like.
 SQ SEQUENCE 320 AA; 35335 MW; 19F14BF58612E08 CRC64;

Query Match 88.2%; Score 1460; DB 2; Length 320;
 Best Local Similarity 86.9%; Pred. No. 3.4e-112;
 Matches 278; Conservative 16; Mismatches 26; Indels 0; Gaps 0;

Qy 1 MPMYOVKHYHGASAPLRYELPTCYRRLPNVSKITSPATDAGVHOENSEPSIOLESROD 60
 Db 1 MPMYOVKHYHGASAPLRYELPTCYRRLPNVSKITSPATDAGVHOENSEPSIOLESROD 60
 Qy 61 DILKRLYEIKAAVDGLSKNITQTPDADLVNTIIQADEPTTLTATNLTDLNSVLGKDYGALK 120

Db 61 DILKRLYEIKAAVDGLSKNITQTPDADLVNTIIQADEPTTLTATNLTDLNSVLGKDYGALK 120
 Qy 121 DIVINANPASPPLSLVLHRLLCERYRVLSTVHTSSVKYVPEMLVKCFGEQARQOSRHE 180
 Db 121 DIVINANPASPPLSLVLHRLLCERYRVLSTVHTSSVKYVPEMLVKCFGEQARQOSRHE 180
 Qy 181 YOLGFTLLWKVVPKTKQMFVSQVTCPIEGEGNINARFLFSLFGQKNAVTLTLIDSWDIA 240
 Db 181 YOLGFTLLWKVVPKTKQMFVSQVTCPIEGEGNINARFLFSLFGQKNAVTLTLIDSWDIA 240
 Qy 241 MFOLREGSSKEKAFAVFSMSNLSALGRSPMLVGNELTVADVLMVSLVQOTGGSSGAAPPTNVQ 300
 Db 241 IFOLKEGSSKEKAFAVFSMSNLSALGRSPMLVGNELTVADVLMVSLVQOTGGSSGAAPPTNVQ 300
 Qy 301 RMLKSCENIAPFPSTALQLLK 320
 Db 301 RMLKSCENIAPFPSTALQLLK 320

RESULT 6
 0983V2 PRELIMINARY; PRT; 280 AA.

AC 0983V2; 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE JTV1.
 GN Name=JTV1;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalek U., Smallos D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
 RA Strausberg R.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC024480; AAH24480.1; -;
 DR MGI; MGI:2385237; JTV1.
 DR InterPro; IPR004046; GST_Cterm.
 DR InterPro; IPR010987; GST_C-like.
 DR Pfam; PF00043; GST_C_1_-like.
 SQ SEQUENCE 280 AA; 31097 MW; E5BB81498983FD2E CRC64;

Query Match 85.3%; Score 1412; DB 2; Length 280;
 Best Local Similarity 100.0%; Pred. No. 2.6e-108;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 QETSEPSIQALESRODDILKRLYEIKAAVDGLSKNITQTPDADLVNTIIQADEPTTLATN 104

Db 5 OETSPSIOALESRODDIKRLYELKAAVDGSKMIHTPDADLDVTNIIQADEPTLATN 64
Qy 105 TLIDNSVYGKXGAKDVIYNNPASPISLIVRLICERYVLTSTHTSSVKNVNP 164
Db 65 TLIDNSVYGKDGAKDVIYNNPASPISLIVRLICERYVLTSTHTSSVKNVNP 124
Qy 165 LVKCGEQAQRKQSRHEVQLGFTLIWKVNVKTKQMFESVQTMCPIDEGENIARFLFSLFGOK 224
Db 125 LVKCGEQAQRKQSRHEVQLGFTLIWKVNVKTKQMFESVQTMCPIDEGENIARFLFSLFGOK 184
Qy 225 HNAVTLTLIDSVWDIAMPOLAREGSSKEKAAPFRSMNSALGRSPMLVGNELTVADVVLMSV 284
Db 185 HNAVTLTLIDSVWDIAMPOLAREGSSKEKAAPFRSMNSALGRSPMLVGNELTVADVVLMSV 244
Qy 285 LOQTGGSGGAAPTINQRMKSCENIAPFSTALQLLK 320
Db 245 LOQTGGSGGAAPTINQRMKSCENIAPFSTALQLLK 280

RESULT 7

Q6DK86 PRELIMINARY; PRT; 311 AA.
AC Q6DK86; 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE MGC9221 protein.
GN Name=MGC69221;
OS Xenopus tropicalis (Western clawed frog) (Sturana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
NCBI_TaxId=8364;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnae.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Ueda T.B., Toshynki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
Krzyniuk M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E.,
Jones S.J., Maira M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [12]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL, BC074561; AAH74561.1; -
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR010987; GST_C_1like.
DR Pfam; PF00043; GST_C_1_1; -
SO SEQUENCE 311 AA; 3480 MW; D98P27F73C466154 CRC64;

Query Match 69.7%; Score 1153.5; DB 2; Length 311;
Best Local Similarity 67.5%; Pred. No. 6,8e-87;
Matches 216; Conservative 49; Mismatches 46; Indels 9; Gaps 3;

Qy 1 MEMYQVKEHGGSABRLVELPTCMYRLPNVHSKTSIPATDAGHVOETSPSIOALESROD 60
Db 1 MEMYQVKEHGGSABRLVELPTCMYRLPNVHSKTSIPATDAGHVOETSPSIOALESROD 52
Qy 61 DILKRLYELKAAVDGSKMIHTPDADLDVTNIIQADEPTLATN TLIDNSVYGKDGAK 120
Db 53 DILKRLYELKAAVDGSKMIHTPDADLDVTNIIQADEPTLATN TLIDNSVYGKDGAK 112
Qy 121 DIVINAPASPISLIVRLICERYVLTSTHTSSVKNVNP LVKCGEQAQRKQSRHE 180
Db 113 DIVINAPASPISLIVRLICERYVLTSTHTSSVKNVNP LVKCGEQAQRKQSRHE 172
Qy 181 YQLGFTLIWKVNVKTKQMFESVQTMCPIDEGENIARFLFSLFGOKHNAVTLTLIDSVWDIA 240
Db 173 YQLGFTLIWKVNVKTKQMFESVQTMCPIDEGENIARFLFSLFGOKHNAVTLTLIDSVWDIA 232
Qy 241 MFOLREGSSKEKAAPFRSMNSALGRSPMLVGNELTVADVVLMSV LOQTGGSGGAAPTINQ 300
Db 233 MFOLREGSSKEKAAPFRSMNSALGRSPMLVGNELTVADVVLMSV LOQTGGSGGAAPTINQ 291
Qy 301 FMLKSCENIAPFSTALQLLK 320
Db 292 FMLKSCENIAPFSTALQLLK 311

RESULT 8

Q6IND4 PRELIMINARY; PRT; 311 AA.
AC Q6IND4; 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE MGC80304 protein.
GN Name=MGC80304;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
NCBI_TaxId=8355;
RN [11]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnae.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Ueda T.B., Toshynki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
Krzyniuk M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E.,
Jones S.J., Maira M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [12]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
Richardson P.;
RL "Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative";
RT Dev. Dyn. 225:384-391(2002).
RN [13]

RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RA Klein S., Gerhard D.S.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC072178; AAH72178.1; -
 DR InterPro; IPR004046; GST_Cterm.
 DR InterPro; IPR010987; GST_C_1like.
 DR Pfam; PF00043; GST_C_1.
 SQ SEQUENCE 311 AA; 34424 MW; D5E8325C18D88751 CRC64;

Query Match
 Best Local Similarity 69.3%; Score 1146.5; DB 2; Length 311;
 Matches 218; Conservative 44; Mismatches 49; Indels 9; Gaps 3;

QY 1 MPMTQVKEPHGSGAPLVELPTCMYRLPNVSKTTPATDAGHVOETSEPSLOALSRQD 60
 DB 1 MPMTQVKEPHGSGAPLVELPTCMYRLPNVSKTTPATDAGHVOETSEPSLOALSRQD 52
 QY 61 DILKRIYELKAAVNGLSKMTHTPDADLVNIIQADEPTLATNTLIDNSVLGKDYGALK 120
 DB 53 DILKRIYELKAAVNGLSKMTHTPDADLVNIIQADEPTLATNTLIDNSVLGKDYGALK 112
 QY 121 DIVINANPASPPLSLVLRHLRCERYVLSVTHSVKVPENLVKRCFQKARXSRHE 180
 DB 113 DIVINANPASPPLSLVLRHLRCERYVLSVTHSVKVPENLVKRCFQKARXSRHE 172
 QY 181 YOLGFTLIMKQVPTQKFSVQTMCPTEGEGNIAFLFSLFGQKHNAAVTLTIDSWVDIA 240
 DB 173 YOLGFTLIMKQVPTQKFSVQTMCPTEGEGNIAFLFSLFGQKHNAAVTLTIDSWVDIA 232
 QY 241 MFOLEGGSSKEKAAPVRSNMSALGRSPMLVGNELTVADVVLMSVLOQTGGSGAAPTVNQ 300
 DB 233 IFOLRBOSSKEKAAPVRSNMSALGRSPMLVGNELTVADVVLMSVLOQTGGSGAAPTVNQ 291
 QY 301 RWLKSCEMLAPFSTALQILK 320
 DB 292 KMKKSCENLASFKSVLRFLK 311

RESULT 9

ID Q7ZYD7 PRELIMINARY; PRT; 311 AA.
 AC Q7ZYD7;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Jv1-prov protein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 NC NCB1_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmitt J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]

RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.,
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative.";
 RL Dev. Dyn. 225:384-391 (2002).
 RN [3]

RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Strausberg R.,
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC043832; AAH43832.1; -
 DR InterPro; IPR004046; GST_Cterm.
 DR InterPro; IPR010987; GST_C_1like.
 DR Pfam; PF00043; GST_C_1.
 SQ SEQUENCE 311 AA; 34647 MW; 2F08C9DA60DD63BE CRC64;

Query Match
 Best Local Similarity 69.3%; Score 1146.5; DB 2; Length 311;
 Matches 218; Conservative 44; Mismatches 49; Indels 9; Gaps 4;

QY 1 MPMTQVKEPHGSGAPLVELPTCMYRLPNVSKTTPATDAGHVOETSEPSLOALSRQD 60
 DB 1 MPMTQVKEPHGSGAPLVELPTCMYRLPNVSKTTPATDAGHVOETSEPSLOALSRQD 52
 QY 61 DILKRIYELKAAVNGLSKMTHTPDADLVNIIQADEPTLATNTLIDNSVLGKDYGALK 120
 DB 53 DILKRIYELKAAVNGLSKMTHTPDADLVNIIQADEPTLATNTLIDNSVLGKDYGALK 112
 QY 121 DIVINANPASPPLSLVLRHLRCERYVLSVTHSVKVPENLVKRCFQKARXSRHE 180
 DB 113 DIVINANPASPPLSLVLRHLRCERYVLSVTHSVKVPENLVKRCFQKARXSRHE 172
 QY 181 YOLGFTLIMKQVPTQKFSVQTMCPTEGEGNIAFLFSLFGQKHNAAVTLTIDSWVDIA 240
 DB 173 YOLGFTLIMKQVPTQKFSVQTMCPTEGEGNIAFLFSLFGQKHNAAVTLTIDSWVDIA 232
 QY 241 MFOLEGGSSKEKAAPVRSNMSALGRSPMLVGNELTVADVVLMSVLOQTGGSGAAPTVNQ 300
 DB 233 IFOLRBOSSKEKAAPVRSNMSALGRSPMLVGNELTVADVVLMSVLOQTGGSGAAPTVNQ 291
 QY 301 RWLKSCEMLAPFSTALQILK 320
 DB 292 KMKKSCENLASFKSVLRFLK 311

RESULT 10

ID Q7T3CO PRELIMINARY; PRT; 321 AA.
 AC Q7T3CO;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE zgc:63976.
 GN ORFNames=zgc:63976;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NC NCB1_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

| Query Match | 56.0%; | Score 926.5; | DB 2; | Length 321; |
|-----------------------|--|---|-----------------------|-------------|
| Best Local Similarity | 57.3%; | Pred. No. 4,1e-68; | | |
| Matches 188; | Conservative 48; | Mismatches 77; | Indels 15; | Gaps 7; |
| QY | 1 | MEMTVQVKEPHYGSGAPRLVRELPTCMYRLNNVHSKTTSS--- | PATDAGHVOETSEPSIQALES | 57 |
| Db | 1 | MPMWQVKEV--SPADITVDLPETCMYKPLPNVHAQSGSLGEHNLQNEV---- | DPTVKALEE | 54 |
| QY | 58 | RODDILKRLYEYLKAAVDGSLSKMHPDADLDVNTNI--LOADEPTLATNTLDLNSVLGK | 114 | |
| Db | 55 | RODEILRLYELKATVDDGLATVTTTPDADLDASTLAHLLTHTTPRAVAVLRGTADLDDLLGK | 114 | |
| QY | 115 | DYGLAKDQIVIANPASPPLSLVLVLRLLLCERYVLSTYHTHSSVKNVPEMLVYKCFG-EOA | 173 | |
| Db | 115 | DLGRLRDIVIANPQPLSLVLVLRLLLCORFOVLSSVHVHSSVSTVPAPLSLCIGPHT | 174 | |
| QY | 174 | RKQSRHEYQLAGFTLIIMKVPKPTOMKFSVOTMCPPIEGEGNIRAFLESLFG-OKHNANVTLT | 232 | |
| Db | 175 | HSYARHRQLGFTLIIMKVDKSLQMKFSTQNMCPPIEGEGNVARFLYRLGLAGEPRDVSATL | 234 | |
| QY | 233 | IDSWVDIAMPOLREGSSHEKAAVFRSSMNSALGRSFWLVGNELTVADVVLMSVLYQOTGSSS | 292 | |
| Db | 235 | MDGVWDVTLFPLAEGSSKERAAVIRALNAAAGRSFWLLGQEFSLADIVSACCVIQTGOTS | 294 | |
| QY | 293 | GAAPTNVQRMILKSCENLAPFSTALQLK | 320 | |
| Db | 295 | -SAPANVQRMILKSCNLGTFSCVDPLQ | 321 | |
| RESULT 11 | | | | |
| ID | Q7Q7A3 | PRELIMINARY; | PRT; | 340 AA. |
| AC | Q7Q7A3; | | | |
| DT | 01-MAR-2004 (TREMBLrel. 26, Created) | | | |
| DT | 01-MAR-2004 (TREMBLrel. 26, Last sequence update) | | | |
| DT | 01-MAR-2004 (TREMBLrel. 26, Last annotation update) | | | |
| DE | AGCP5808 (Fragment) | | | |
| GN | Name=agCG50514; ORFNames=ENSANGG0000011827; | | | |
| OS | Anopheles gambiae str. PEST. | | | |
| OC | Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; | | | |
| CC | Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles. | | | |
| CX | NCBI_TaxId=160454; | | | |
| NP | (1) | | | |
| SEQUENCE FROM N.A. | | | | |

| | | | | |
|-----------------------|-------|--------------|-------|-------------------------------------|
| Query Match | 14.9% | Score 247 | DB 2 | Length 340 |
| Best Local Similarity | 28.7% | Pred. No. 6 | 16-12 | |
| Matches | 94 | Conservative | 49 | Mismatches 118; Indels 66; Gaps 13. |

| | | | |
|----|-----|--|-----|
| QY | 3 | MYQVKEPHYGGSAAPRLVLEPTCMYRLPNVYHSKTTSP-----ATD---AGHV----- | 44 |
| Db | 12 | MYRLKAPVM---NQDVCELPCTCMYTLKPVCAVSHDPKAFSADSATGALAAAGPVDMSIEID | 69 |
| QY | 45 | ---QETSEPLQLAESQDDILIKRLYLKKAAVDGLSKMIRTPDDLDV--TIIIDADEFTT | 100 |
| Db | 70 | NLLQKAENDELRMLAEKQORVYLOQLAEIK-----KEINAMRTKLTNNAPAPVQGST | 122 |
| QY | 101 | -----LATNTDINSVLGKDYGALADIYINANPASRPSTLYLHRLTLCERYVLSTVH | 153 |
| Db | 123 | PLSKSAQLKAEPIWLT-----CLDDPVMNASPEYVPSYSLAKLIMKDRMLQVCECF | 174 |
| QY | 154 | THSSVKNVPE-----NLVKCFGEARKQSHREYQLGFTLIWKVY-PTYQMKFSVQTMCP | 206 |
| Db | 175 | THSTVPRKLSEBALAFQNAVTAASGTAAANLPR-----IKVTILMKVGAVTETIITSPTSVP | 230 |
| QY | 207 | IEEGGNIAAPLFSL-----FGQKNAVTLTLIDSWVDIAMFQLRBEGSSKEKAAVFPMN | 260 |
| Db | 231 | ICGEVNILRYLIRSCGPEEFNVEQDDN---VDEVDISILDACYLLINKONVAKROOILRTLG | 287 |
| QY | 261 | SALGRSPWLYGNELTVADVVLMSYLOQ | 287 |
| Db | 288 | AKLGKAAGFGAGADISTLCDIAFTSAVKQ | 314 |

| RESULT 12 | 07KJMS | PRELIMINARY; | PRT; | 301 AA. |
|-----------|--|--------------|------|---------|
| ID | 07KJMS | | | |
| AC | 07KJMS; | | | |
| DT | 05-JUL-2004 (TREMBlrel, 27, Created) | | | |
| DT | 05-JUL-2004 (TREMBlrel, 27, Last sequence update) | | | |
| DT | 05-JUL-2004 (TREMBlrel, 27, Last annotation update) | | | |
| DE | CG12304-BB. | | | |
| GN | ORFNames=CG12304; | | | |
| OS | Drosophila melanogaster (Fruit fly). | | | |
| OC | Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; | | | |
| OC | Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; | | | |
| OC | Ephydroidea; Drosophilidae; Drosophila. | | | |
| OX | NCBI_TaxID=7227; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEBLIN=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185; | | | |
| RA | Adams D.C., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D., | | | |
| RA | Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., | | | |
| RA | George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., | | | |
| RA | Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., | | | |
| RA | Brandon R.C., Rogers Y.H., Blazel R.G., Champe M., Pfeiffer B.D., | | | |
| RA | Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L., | | | |
| RA | Abvill J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D., | | | |
| RA | Ballew R.M., Baau A.B., Baxendale J., Bayraktaroglu L., Beasley E.M., | | | |
| RA | Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., | | | |
| RA | Botkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P., | | | |
| RA | Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., | | | |
| RA | Cheriy J.M., Cawley S., Dahlke C., Davidson L.B., Davies P., | | | |
| RA | de Pallos B., Delcher A., Deng Z., May A.D., Dew I., Dietz S.M., | | | |
| RA | Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., | | | |
| RA | Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., | | | |
| RA | Foster C., Gabrielson A.E., Garg N.S., Gelbart W.M., Glasser K., | | | |
| RA | Gloder A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., | | | |

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Houston D., Houston K.A., Howland T.J., Mei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merklov G., Mlshina N.V., Moberly G., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuseken D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Furt V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Sidenklamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weinstein J.,
 RA Williams S.W., Woodgerf, Morley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan W., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster.",
 RL Science 287:2185-2195 (2000).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodegren E.J.,
 RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
 RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
 RT melanogaster euchromatic genome sequence.",
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.,
 RT "The transposable elements of the Drosophila melanogaster euchromatin:
 RT a genomics perspective.",
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.,
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review.",
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 [5]
 RP SEQUENCE FROM N.A.
 RG FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RG FlyBase;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE005330; AAS6498.1; -.
 DR InterPro; IPR004046; GST_Cterm.
 DR InterPro; IPR010987; GST_C_1like.
 DR Pfam; PF00043; GST_C_1.
 SQ SEQUENCE 301 AA; 33241 MW; 35224E17FE03F3BE CRC64;

Query Match 13.5%; Score 224; DB 2; Length 301;
 Best Local Similarity 28.6%; Pred. No. 4.1e-10;
 Matches 86; Conservative 47; Mismatches 122; Indels 46; Gaps 9;

QY 3 MYGVKPYHGGASAPLRYELPTCYRLPNV-----HSKTTSPATDAGHVQETSPSLQA 54
 DB 1 MYELKTL---LPQFDIKLPTCYPLKVNSLADSLASGSSISASTSA---STSSCDPTAS 54
 QY 55 LBSRODDILKRLYELKAAVDGSKNIHPDDLDVTNTLQADDEPTLATNTLDNSVLGK 114
 DB 55 VAARQEKVLKQLELKAQGLQGR-----NGLVCG--KTFQHTTAFONG-----GL 98
 QY 115 DYGLAKDIYINANPASPPLSLVLRLLCERYRVLSVHTHSSVQNV-----PENLVK 167
 DB 99 KEVPLQDVVINGHPNFIYPALLALKNAMRNLYTIDVKTFTHTSMADIGPAREFEANLAK 158
 QY 168 CGEGARQKRSREYVGLGFTLLTKRVNPKQMKFSVQTMCPPIGEGNIARFLSLFGQRNA 227
 DB 159 VPVNPALP-----KISVTLTWKCEHTEMTSSPTMVPVIGEVNIIIRYLGRVGPAREYR 212
 QY 228 VTLTL---IDSWVDIAMPQLREGSKCAAVFRSNASLAGSPVLVGNELVAVLVMSV 284
 DB 213 EGSPLCNEIDLVDICYGLACNHTKQVAMVRLIDRLKQKQYFGSGSQMSVADVYSS 272
 QY 285 L 285
 DB 273 L 273
 RESULT 13
 ID 08T060 PRELIMINARY; PRT; 313 AA.
 AC 08T060;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
 DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 LN LD25772P (Fragment).
 GN ORNames=CG12304;
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunco J., Pacleb J., Farasas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.,
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY069537; ALJ39682.2; -.
 DR FlyBase; FBgn0036515; CG12304.
 DR InterPro; IPR004046; GST_Cterm.
 DR InterPro; IPR010987; GST_C_1like.
 DR Pfam; PF00043; GST_C_1.
 FT NON_TER 1
 SQ SEQUENCE 313 AA; 34522 MW; 5AF9C801877F8531 CRC64;

Query Match 13.5%; Score 224; DB 2; Length 313;
 Best Local Similarity 28.6%; Pred. No. 4.3e-10;
 Matches 86; Conservative 47; Mismatches 122; Indels 46; Gaps 9;

QY 3 MYGVKPYHGGASAPLRYELPTCYRLPNV-----HSKTTSPATDAGHVQETSPSLQA 54
 DB 13 MYELKTL---LPQFDIKLPTCYPLKVNSLADSLASGSSISASTSA---STSSCDPTAS 66
 QY 55 LBSRODDILKRLYELKAAVDGSKNIHPDDLDVTNTLQADDEPTLATNTLDNSVLGK 114
 DB 67 VAARQEKVLKQLELKAQGLQGR-----NGLVCG--KTFQHTTAFONG-----GL 110
 QY 115 DYGLAKDIYINANPASPPLSLVLRLLCERYRVLSVHTHSSVQNV-----PENLVK 167
 DB 111 KEVPLQDVVINGHPNFIYPALLALKNAMRNLYTIDVKTFTHTSMADIGPAREFEANLAK 170

```
OY 168 CFGEQARKOSRHEYOGLFTLWKNVPTKOTMFKFSVOTMCIEEGNIAFLFSLFGQKHA 227
DB 171 VFNVALP-----KISVTLWKNCHENTMISPTMYVPIYGEVNIJRIIGVGAPAEHY 224
OY 228 VTLTL---IDSWVDIAMFOLREGSSKEKAVFRSMNSALGRSPWLVGNELTVADVVLMV 284
DB 225 EGSPLCNEIDLVDICYQLLRGNTHTKTQVAMVRLDLKRLKQOXYFGKISQMSVADYGVSS 284
OY 285 L 285
DB 285 L 285

RESULT 14
O6NMK4 PRELIMINARY; PRT; 322 AA.
ID O6NMK4
AC O6NMK4;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE LP12114p (Fragment).
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Carlson J., Chavez C., Friese E., George R., Pacleb J.,
RA Park S., Wan K., Yu C., Rubin G.M., Celniker S.;
RA Submitted (May-2004) to the EMBL/Genbank/DBJ databases.
DR EMBL; BT012670; AAT00476.1; -.
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR010987; GST_C-like.
DR Pfam; PF00043; GST_C_1.
FT NON TER
SQ SEQUENCE 322 AA; 35528 MW; 82A15927DF928DAE CRC64;

Query Match 13.5%; Score 224; DB 2; Length 322;
Best Local Similarity 28.6%; Pred. No. 4,5e-10;
Matches 86; Conservative 47; Mismatches 122; Indels 46; Gaps 9;

OY 3 MYQVPEYHGSAPLVELPTCMYRLPNV-----HSTTSPADNAGVQTSSESLQA 54
DB 22 MYELKTL--LPQFDIKLPTCMYRLPNVSLADSLASGSSTASISA--STSSCDRTAS 75
OY 55 LESRODDILKRLYELKAAVDGLSKMHTPDADLDVTNIIQADBPPTLATNTLDLSLVGK 114
DB 76 VAAREKVKUQLELKAQUGQR-----AGLGUG--KTQHTTAPONG-----GL 119
OY 115 DYGALKDIVINANPASPRLSLVLRHLLCERYRVLTSTHTSSVKNV-----PENLVK 167
DB 120 KEVPLDQVYINGNHPIFYALLALKNAMRNLYTIDVKTFTHTMDIGPARAREFEANLAK 179
OY 168 CEGEQARKOSRHEYOGLFTLWKNVPTKOTMFKFSVOTMCIEEGNIAFLFSLFGQKHA 227
DB 180 VFNVALP-----KISVTLWKNCHENTMISPTMYVPIYGEVNIIRIIGVGAPAEHY 233
OY 228 VTLTL---IDSWVDIAMFOLREGSSKEKAVFRSMNSALGRSPWLVGNELTVADVVLMV 284
DB 224 EGSPLCNEIDLVDICYQLLRGNTHTKTQVAMVRLDLKRLKQOXYFGKISQMSVADYGVSS 293
OY 285 L 285
DB 294 L 294

RESULT 15
MC02_DROME STANDARD; PRT; 334 AA.
ID MC02_DROME
AC 09VUR3;
DT 16-OCT-2001 (Rel. 40, Created)
```

```
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Probable multisynthetase complex auxiliary component p38.
DB ORNames=CG12304;
GN ORFNames=CG12304;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galante R.F.,
RA Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
RA Suton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benoe P.V., Berman B.P., Bhattacharya D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hootin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Talali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levytsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mikhlin N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacleb J.M.,
RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
RA Reinert C., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Switek R., Tector C., Turner C., Turner E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP GENOME REANNOTATION.
RX MEDLINE=22426069; PubMed=12537572;
RA Miura S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Battencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review."
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
CC -!- FUNCTION: Probable core protein of the multisynthetase complex
CC that serves as a template for the assembly of the supramolecular
CC structure (By similarity).
CC -!- SUBUNIT: Component of the multisynthetase complex which is
CC comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the
CC monospecific isoleucyl, leucyl, glutamyl, methionyl, Iysyl,
CC arginyl, and aspartyl-tRNA synthetases as well as three auxiliary
CC proteins, p18, p48 and p43 (By similarity).
CC -!- SIMILARITY: Contains 1 GST-like domain.
```

```

-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.ebi.ch/announce/
CC      or send an email to license@ebi-sib.ch).
CC      -----
DR      EMBL; AB003530; AAF49612.1; -.
DR      IntAc; Q9VUR3; -.
DR      FlyBase; FBgn0036515; CG12304.
DR      InterPro; IPR010987; GST_C like.
DR      InterPro; IPR004046; GST_C term.
DR      Pfam; PF00043; GST_C; 1.
KM      Protein bioynthesis.
FT      DOMAIN 280 327 GST-like.
SQ      SEQUENCE 334 AA; 36933 MW; B68FD70AE621990F CRC64;

Query Match      13.1%; Score 216.5; DB 1; Length 334;
Best Local Similarity 26.0%; Pred. No. 2e-09;
Matches 86; Conservative 48; Mismatches 124; Indels 73; Gaps 9;

QY      3 MYQVKPYHGGSAFLRVLPCTMTRLPNV-----HSKTTSPATPDAG----- 42
      1 MYELKTL--LPOFDIKLPTCMYPLKQVSLAADSLASGSSTASSTSSCKLEANRID 57
QY      43 -----HVQETSEPSLQALSESRODDILKRLYEELKAVDGLSKMHTPD 84
      58 RTGRNATCTALDLSLGRQIQRLIKODDTASVAAKQERVLKQLELKAQLGQIR----- 110
QY      85 ADLDVTNIIQADEPPTLATNTLDLNSVLGKQYALNDIVINANPASPLSLVLRHLC 144
      111 AGLGVCG--KTFQHTTAFQNG-----GLKEVPLQDVINGHNPFIYALLALKNAMRN 161
QY      145 RYRVLSTVTHHSSVKNV-----PENLYKCFGEQARKOSRHEVQLGFTLIMKNVPTQM 197
      162 LYTIIDVKTFTHSTMADIGPAAREFEANLAKVPVNPALP-----KISVTLIMKNCETEM 215
QY      198 KFSVQTMCPLEGEGNIARFLFSLFGQKHNAVTLTL--IDSWVDIAMFQLRGSSSEKAA 254
      216 ISSEPIMVYPIYGEVNIIRYLGRVGAPEYRREGSPLCNEIDLVDICYQLRCNTHKTQYA 275
QY      255 VFRSMNSALGRSPWLVGNEELTVADVVLMSVL 285
      276 MVRLLDKRLQKQYFEGSQMSVADVGVYSSL 306
DB

```

Search completed: February 23, 2005, 13:57:15
 Job time : 61.9811 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:25:39 ; Search time 102.258 Seconds
(without alignment)
1758.725 Million cell updates/sec

Title: US-10-622-817-8

Perfect score: 2596

Sequence: 1 MIVFVRFNSSHGFPVEVSDS.....CWNCGCEMNRVCMGDHMDV 465

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq.16Dec04:*

1: geneseqp19808:*\n2: geneseqp19908:*\n3: geneseqp20008:*\n4: geneseqp20018:*\n5: geneseqp20028:*\n6: geneseqp20038:*\n7: geneseqp20038:*\n8: geneseqp20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|---------------------|
| 1 | 2596 | 100.0 | 465 | 2 | AAV32501 Human par |
| 2 | 2596 | 100.0 | 465 | 6 | AAE30800 Human par |
| 3 | 2586 | 100.0 | 465 | 6 | ABO07156 Human p53 |
| 4 | 2584 | 99.5 | 465 | 6 | AAE30801 Human par |
| 5 | 2405 | 92.6 | 437 | 2 | AAV32502 Human par |
| 6 | 2405 | 92.6 | 437 | 6 | ABO07157 Human p53 |
| 7 | 2170.5 | 83.6 | 464 | 4 | AAAB67517 Amino aci |
| 8 | 2165.5 | 83.4 | 464 | 4 | AAAB67531 Amino aci |
| 9 | 2165.5 | 83.4 | 464 | 4 | AAAB67532 Amino aci |
| 10 | 2087.5 | 80.4 | 451 | 4 | AAAB67533 Amino aci |
| 11 | 1696.5 | 65.4 | 316 | 6 | ABO07158 Human p53 |
| 12 | 1679.5 | 64.7 | 344 | 4 | AAAB67521 Amino aci |
| 13 | 1234.5 | 47.6 | 296 | 4 | AAAB67526 Amino aci |
| 14 | 1058.5 | 40.8 | 468 | 4 | ABBE5114 Drosophila |
| 15 | 1002 | 38.6 | 250 | 4 | AAAB67519 Amino aci |
| 16 | 997 | 38.4 | 262 | 4 | AAAB67518 Amino aci |
| 17 | 870 | 33.5 | 156 | 4 | AAAB67525 Amino aci |
| 18 | 671.5 | 25.9 | 183 | 4 | AAAB67524 Amino aci |
| 19 | 671.5 | 25.9 | 194 | 4 | AAAB67524 Amino aci |
| 20 | 544.5 | 21.0 | 386 | 4 | AAAB6951 C. elegans |
| 21 | 306 | 19.5 | 153 | 4 | AAAB67523 Amino aci |
| 22 | 335 | 12.9 | 77 | 4 | AAAB67529 Amino aci |
| 23 | 263 | 10.1 | 46 | 4 | AAAB17055 Peptide # |
| 24 | 263 | 10.1 | 46 | 4 | ABB36052 Peptide # |
| 25 | 263 | 10.1 | 46 | 4 | AAAB29548 Peptide # |

| | | | | | |
|----|-------|------|-----|---|----------------------|
| 26 | 263 | 10.1 | 46 | 4 | ABB30873 Peptide # |
| 27 | 263 | 10.1 | 46 | 4 | ABB21452 Protein # |
| 28 | 263 | 10.1 | 46 | 4 | AAAB69222 Human bon |
| 29 | 263 | 10.1 | 46 | 4 | AAAB69222 Human bon |
| 30 | 263 | 10.1 | 46 | 4 | AAAB69222 Human bon |
| 31 | 263 | 10.1 | 46 | 4 | AAAB69222 Human bon |
| 32 | 252 | 9.7 | 106 | 8 | ADMG6489 Modified |
| 33 | 251 | 9.7 | 63 | 4 | AAAB67522 Amino aci |
| 34 | 251 | 9.7 | 105 | 4 | AAAB67520 Amino aci |
| 35 | 251 | 9.7 | 105 | 8 | ADMG6487 Modified |
| 36 | 240.5 | 9.3 | 503 | 4 | ABBE1708 Drosophila |
| 37 | 240.5 | 9.3 | 503 | 7 | ADBE7082 Atrialade-1 |
| 38 | 240 | 9.2 | 511 | 4 | ABBE63665 Drosophila |
| 39 | 239 | 9.2 | 520 | 8 | ADBE22562 Sea-squid |
| 40 | 236.5 | 9.1 | 445 | 3 | AAAB21034 Human nuc |
| 41 | 236.5 | 9.1 | 445 | 8 | ADQ96154 T cell ac |
| 42 | 236.5 | 9.1 | 445 | 8 | ADQ96026 T cell ac |
| 43 | 236.5 | 9.1 | 557 | 3 | AAAB98059 Human Rln |
| 44 | 236.5 | 9.1 | 557 | 8 | ADQ96028 T cell ac |
| 45 | 236 | 9.1 | 53 | 4 | AAAB67528 Amino aci |

ALIGNMENTS

RESULT 1
ID AAY32501 standard; protein; 465 AA.
AC AAY32501;
XX
XX
XX 21-OCT-1999 (first entry)
XX
XX Human parkin gene variant protein.
XX
XX Parkinson's disease related gene; parkin gene; variant; gene therapy.
XX
XX Homo sapiens.
XX
XX MO9940191-A1.
XX
XX 12-AUG-1999.
XX
XX 09-FEB-1999; 99WC-JP000545.
XX
XX 09-FEB-1998; 98JP-00027531.
XX
XX (SHIM/) SHIMIZU N.
XX (MIZU/) MIZUNO Y.
XX
XX Shimizu N, Mizuno Y;
XX
XX WPI; 1999-494295/41.
XX DR N-PSDB; AAX99923.
XX
XX Gene implicated in the pathology of Parkinson's disease, used for
XX PT treatment of the disease.
XX
XX
XX Claim 1; Page 83-88; 114pp; English.
XX
XX
XX This sequence is encoded by a gene of the invention, and is implicated in
XX the pathology of Parkinson's disease. This sequence is a variant of the
XX parkin gene found in parkinson's disease patients. The sequences may be
XX used for the diagnosis, treatment (including gene therapy) and
XX investigation of Parkinson's disease
XX
XX Sequence 465 AA;
XX
XX Query Match 100.0%; Score 2596; DB 2; Length 465;
XX Best Local Similarity 100.0%; Pred. No. 2.3e-219; Indels 0; Gaps 0;
XX Matches 465; Conservativeness 0; Mismatches 0;
XX
XX 1 MIVFVRFNSSHGFPVEVSDTSIFOLKEVAVAKRGVPAQDLRVIFAGKELRNDWTQVNCND 60

```
Db 1 MIVFVRFNSSHGPFVEVSDTSIFQLKEVYAKRGVPAQDLVIRFAGKELRNDMTVQNC 60
Qy 61 LDOQSIVHIVORPMWKGEMNATGDDPRNNAAGCEREPOSILTRVDLSSVLPEDSVGLA 120
Db 61 LDOQSIVHIVORPMWKGEMNATGDDPRNNAAGCEREPOSILTRVDLSSVLPEDSVGLA 120
Qy 121 VILHTDSRKDSPPAGSPAGRSIYNSFYVCKGPCORVOPGKLRYOCSTCRQATLTLTQGP 180
Db 121 VILHTDSRKDSPPAGSPAGRSIYNSFYVCKGPCORVOPGKLRYOCSTCRQATLTLTQGP 180
Qy 181 SCMDVVLIPNMSGECQSPHCPTSAEFPFKCGAHPSTDKETPVLAHLIATNSRNTICIT 240
Db 181 SCMDVVLIPNMSGECQSPHCPTSAEFPFKCGAHPSTDKETPVLAHLIATNSRNTICIT 240
Qy 241 CTDVRSPLYVFOCNSRHVYICLDCPHLYCVTRLNDQFVHDQGLGSLPCVAGCNSLILKE 300
Db 241 CTDVRSPLYVFOCNSRHVYICLDCPHLYCVTRLNDQFVHDQGLGSLPCVAGCNSLILKE 300
Qy 301 LHHFRILGEBQYNYRYOYGABECVYLQMGVLCPRPGCAGLLPBPDRKVTCEGNGLGC 360
Db 301 LHHFRILGEBQYNYRYOYGABECVYLQMGVLCPRPGCAGLLPBPDRKVTCEGNGLGC 360
Qy 361 GFAPCRECKEAYHGECSAVFEASGTTTQAYRVDERAABEQARWEAASKETIKKTTKPCR 420
Db 361 GFAPCRECKEAYHGECSAVFEASGTTTQAYRVDERAABEQARWEAASKETIKKTTKPCR 420
Qy 421 CHVPEKXGCMHMKCPQPCRLEMCNCGCENNRVCMGDHWFV 465
Db 421 CHVPEKXGCMHMKCPQPCRLEMCNCGCENNRVCMGDHWFV 465
```

RESULT 2

```
AAE30800
ID AAE30800 standard; protein; 465 AA.
AC AAE30800;
DT 24-FEB-2003 (first entry)
XX
DE Human Parkin protein.
XX
KW Human; Parkin protein; neurological disorder; apoptosis; gene therapy;
KW ischaemic stroke; Parkinson's disease; Alzheimer's disease; noctropic;
KW transgenic; cerebroprotective; neuroprotective; neurotransplantation.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Cleavage-site 126
PN MO200279459-A2.
XX
PD 10-OCT-2002.
XX
PE 02-APR-2002; 2002WC-DK000221.
XX
PR 29-MAR-2001; 2001DK-00000525.
PR 03-APR-2001; 2001US-0281286P.
XX
PA (NSGE-) NSGENE AS.
PI Jensen PH;
XX
DR WPI; 2003-046812/04.
DR N-PSDB; AAD47679.
XX
PT New isolated nucleic acid sequence encoding a Parkin polypeptide, useful
PT for treating, preventing or diagnosing neurological disorders, e.g.
PT Parkinson's disease, Alzheimer's disease or ischemic stroke, and in
PT screening assays.
XX
PS Claim 10; Page 69; 71pp; English.
```

```
XX
CC The invention relates to Parkin protein and its corresponding nucleic
CC acid sequence. The nucleic acid sequence is useful for altering the
CC proteolytic processing of Parkin at its potential cleavage site at Asp
CC 126. The invention is used in manufacturing or testing a pharmaceutical
CC composition for treating and/or preventing a neurological disorder, e.g.
CC Alzheimer's disease or ischaemic stroke. It also used for detecting the
CC occurrence of proteolytic processing of Parkin at Asp 126 in a sample, in
CC monitoring a potential disposition for a neurodegenerative disease and
CC for treating, preventing and/or diagnosing Parkinson's disease or other
CC neurodegenerative disorders. The viral vector is used for transforming
CC neuronal cells in vivo or ex vivo. The invention is useful for
CC neurotransplantation into the CNS of a mammal. It may be used in
CC screening assays to identify compounds that increase or decrease
CC apoptosis. It is also used in gene therapy. The present sequence is human
CC Parkin protein
CC
SQ Sequence 465 AA;
```

Query Match 100.0%; Score 2596; DB 6; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.3e-219;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 MIVFVRFNSSHGPFVEVSDTSIFQLKEVYAKRGVPAQDLVIRFAGKELRNDMTVQNC 60
Db 1 MIVFVRFNSSHGPFVEVSDTSIFQLKEVYAKRGVPAQDLVIRFAGKELRNDMTVQNC 60
Qy 61 LDOQSIVHIVORPMWKGEMNATGDDPRNNAAGCEREPOSILTRVDLSSVLPEDSVGLA 120
Db 61 LDOQSIVHIVORPMWKGEMNATGDDPRNNAAGCEREPOSILTRVDLSSVLPEDSVGLA 120
Qy 121 VILHTDSRKDSPPAGSPAGRSIYNSFYVCKGPCORVOPGKLRYOCSTCRQATLTLTQGP 180
Db 121 VILHTDSRKDSPPAGSPAGRSIYNSFYVCKGPCORVOPGKLRYOCSTCRQATLTLTQGP 180
Qy 181 SCMDVVLIPNMSGECQSPHCPTSAEFPFKCGAHPSTDKETPVLAHLIATNSRNTICIT 240
Db 181 SCMDVVLIPNMSGECQSPHCPTSAEFPFKCGAHPSTDKETPVLAHLIATNSRNTICIT 240
Qy 241 CTDVRSPLYVFOCNSRHVYICLDCPHLYCVTRLNDQFVHDQGLGSLPCVAGCNSLILKE 300
Db 241 CTDVRSPLYVFOCNSRHVYICLDCPHLYCVTRLNDQFVHDQGLGSLPCVAGCNSLILKE 300
Qy 301 LHHFRILGEBQYNYRYOYGABECVYLQMGVLCPRPGCAGLLPBPDRKVTCEGNGLGC 360
Db 301 LHHFRILGEBQYNYRYOYGABECVYLQMGVLCPRPGCAGLLPBPDRKVTCEGNGLGC 360
Qy 361 GFAPCRECKEAYHGECSAVFEASGTTTQAYRVDERAABEQARWEAASKETIKKTTKPCR 420
Db 361 GFAPCRECKEAYHGECSAVFEASGTTTQAYRVDERAABEQARWEAASKETIKKTTKPCR 420
Qy 421 CHVPEKXGCMHMKCPQPCRLEMCNCGCENNRVCMGDHWFV 465
Db 421 CHVPEKXGCMHMKCPQPCRLEMCNCGCENNRVCMGDHWFV 465
```

RESULT 3

```
ABO07156
ID ABO07156 standard; protein; 465 AA.
AC ABO07156;
DT 13-AUG-2003 (first entry)
XX
DE Human p53 modifying protein, SEQ ID 116.
XX
KW Human; p53 modifier; cyrostatic; cancer; cyrostatic; antiangiogenic;
KW antiapoptotic; p53 pathway; breast cancer; colon cancer; kidney cancer;
KW lung cancer; ovarian cancer; angiogenesis; cell cycle;
KW apoptotic disorder; cell proliferation disorder.
XX
OS Homo sapiens.
XX
```

XX W0200299122-A1.
 XX 12-DEC-2002.
 PD
 PP 03-JUN-2002; 2002WO-US017382.
 XX
 PR 05-JUN-2001; 2001US-0296076P.
 PR 10-OCT-2001; 2001US-0328605P.
 PR 15-FEB-2002; 2002US-0357253P.
 XX
 XX (EXEL-) EXELIXIS INC.
 PA
 PI Friedman L, Plozman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
 PI WPI, 2003-156859/15.
 DR N-PSDB; ACD13332.
 XX
 PT Identifying modulators of the p53 pathway for use in treating apoptotic
 PT or cell proliferation disorders, comprises screening for agents that
 PT modulate activity of a human ortholog of genes that modify the p53
 PT pathway in Drosophila.
 XX
 PS Example 2; Page 399-401; 678bp; English.
 XX
 CC The invention relates to identifying (M1) a candidate p53 pathway
 CC modulating agent. By contacting an assay system comprising a purified HM
 CC polypeptide (human orthologue of genes that modify the p53 pathway in
 CC Drosophila) or nucleic acid with a test agent under conditions, where but
 CC for the presence of the test agent, the system provides a reference
 CC activity, and detecting a test agent-biased activity of the assay system.
 CC Also included are modulating (M2) a p53 pathway of a cell (comprising
 CC contacting a cell defective in p53 function with a candidate modulator
 CC that specifically binds to a HM polypeptide comprising an HM amino acid
 CC sequence, where p53 function is restored), modulating (M3) a p53 pathway
 CC in a mammalian cell (comprising contacting the cell with an agent that
 CC specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
 CC a disease in a patient (comprising: (a) obtaining a biological sample
 CC from the patient; (b) contacting the sample with a probe for HM
 CC expression; (c) comparing the results with a control; and (d) determining
 CC whether the comparison indicates a likelihood disease). (M1) is useful
 CC for identifying modulators of the p53 pathway. A probe for HM expression
 CC is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,
 CC in a patient, where the cancer has greater than 25 % expression level.
 CC Modulators identified by (M1) are useful in a variety of diagnostic and
 CC therapeutic applications, where disease or disorder prognosis is related
 CC to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
 CC proliferation disorders (e.g. cancer). Another two new methods (M2 and
 CC M3) are useful for modulating the p53 pathway of a cell, thus restoring
 CC the p53 function of the cell, so that the cell undergoes normal
 CC proliferation or progression through the cell cycle. (M2) and (M3) are
 CC also useful for treating defects in the p53 pathway such as angiogenic,
 CC apoptotic or cell proliferation disorders. The present sequence
 CC represents a human p53 pathway modifying protein
 CC
 XX
 XX Sequence 465 AA:
 SQ
 Query Match 100.0%; Score 2596; DB 6; Length 465;
 Best Local Similarity 100.0%; Pred. No. 2,3e-219;
 Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----------|--|---|-----|
| Db | 181 | SCMDVDLIPRNMGEQCSHPCEPSTAEFFPKGAHPTSDKEFVALHLATNSRNTCTT | 24 |
| Qy | 241 | CTDVRSPVLVFGQNSRHVTCIDCFHLVCTRLNDROPVHDPOLGSLPCVACGPNLSLIXE | 30 |
| Db | 241 | CTDVRSPVLVFGQNSRHVTCIDCFHLVCTRLNDROPVHDPOLGSLPCVACGPNLSLIXE | 30 |
| Qy | 301 | LHHFRILGEEQVNRYYQYGAEECVLMGGVLCRRPGCGAGLLPEPDRKVTCEGSGNLGC | 36 |
| Db | 301 | LHHFRILGEEQVNRYYQYGAEECVLMGGVLCRRPGCGAGLLPEPDRKVTCEGSGNLGC | 36 |
| Qy | 361 | GPAFCRECKAYHEGECSAVFEASGTTTQAYRVDEPAARAEARAEASKETIKKTKPCPR | 42 |
| Db | 361 | GPAFCRECKAYHEGECSAVFEASGTTTQAYRVDEPAARAEARAEASKETIKKTKPCPR | 42 |
| Qy | 421 | CHVPEKNQGMHMKCPDPQCRLEWCMNCGCEMNRVCMGDHWEV | 465 |
| Db | 421 | CHVPEKNQGMHMKCPDPQCRLEWCMNCGCEMNRVCMGDHWEV | 465 |
| RESULT 4 | | | |
| AAE30801 | | | |
| ID | AAE30801 | standard; protein; 465 AA. | |
| XX | AAE30801; | | |
| AC | AAE30801; | | |
| XX | | | |
| DT | 24-FEB-2003 | (first entry) | |
| XX | | | |
| DE | Human Parkin D126 mutant protein. | | |
| XX | | | |
| KW | Human; Parkin protein; neurological disorder; apoptosis; gene therapy; | | |
| KW | ischemic stroke; Parkinson's disease; Alzheimer's disease; neurotropic; | | |
| KW | transgenic; cerebroprotective; neuroprotective; neurotransplantation; | | |
| KW | mutant; mutein. | | |
| XX | | | |
| OS | Homo sapiens. | | |
| OS | Synthetic. | | |
| XX | | | |
| XX | Key | Location/Qualifiers | |
| FT | Misc-difference 126 | | |
| FT | /note= "Wild-type Asp is replaced with Glu" | | |
| FT | Cleavage-site 126 | | |
| FT | Misc-difference 223 | | |
| FT | /note= "Encoded by CCA" | | |
| XX | | | |
| XX | WO200279459-A2. | | |
| XX | | | |
| PD | 10-OCT-2002. | | |
| XX | | | |
| XX | 02-APR-2002; 2002WO-DK000221. | | |
| XX | | | |
| XX | 29-MAR-2001; 2001DK-00000525. | | |
| PR | 03-APR-2001; 2001US-0281286P. | | |
| XX | | | |
| XX | (NSGE-) NSGENE AS. | | |
| XX | | | |
| XX | Jensen PH; | | |
| XX | | | |
| XX | WPI: 2003-046812/04. | | |
| DR | N-PSDB; AAD47680. | | |
| XX | | | |
| XX | New isolated nucleic acid sequence encoding a Parkin polypeptide, useful | | |
| PT | for treating, preventing or diagnosing neurological disorders, e.g. | | |
| PT | Parkinson's disease, Alzheimer's disease or ischemic stroke, and in | | |
| PT | screening assays. | | |
| XX | | | |
| XX | Claim 10; Page 71; 71pp; English. | | |
| XX | | | |
| CC | The invention relates to Parkin protein and its corresponding nucleic | | |
| CC | acid sequence. The nucleic acid sequence is useful for altering the | | |
| CC | proteolytic processing of Parkin at its potential cleavage site at Asp | | |
| CC | 126. The invention is used in manufacturing or testing a pharmaceutical | | |
| CC | composition for treating and/or preventing a neurological disorder, e.g. | | |

CC Alzheimer's disease or ischaemic stroke. It also used for detecting the
CC occurrence of proteolytic processing of parkin at Asp 126 in a sample, in
CC monitoring a potential disposition for a neurodegenerative disease, and
CC for treating, preventing and/or diagnosing Parkinson's disease or other
CC neurodegenerative disorders. The viral vector is used for transforming
CC neuronal cells in vivo or ex vivo. The invention is useful for
CC neurotransplantation to identify the CNS of a mammal. It may be used in
CC screening assays to identify compounds that increase or decrease in
CC apoptosis. It is also used in gene therapy. The present sequence is human
CC parkin mutant protein
XX

SQ Sequence 465 AA;

Query Match 99.5%; Score 2584; DB 6; Length 465;
Best Local Similarity 99.6%; Pred. No. 2.6e-218;

Matches 463; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MIVFRRFNSHGFPVEVSDTISIFOLKEVAKRGQVADQRLVIFAGKELRNDMTVQNC 60
DB 1 MIVFRRFNSHGFPVEVSDTISIFOLKEVAKRGQVADQRLVIFAGKELRNDMTVQNC 60
QY 61 LDQOSIVHIVORPMRKGEENATGDDPRNAAAGCEREPOSITRVLDSSSVLPEDSVGLA 120
DB 61 LDQOSIVHIVORPMRKGEENATGDDPRNAAAGCEREPOSITRVLDSSSVLPEDSVGLA 120
QY 121 VILHTDSRKDSPPGSPAGRSINYSFYVYCKGPCQORVOPGKLRVQGSTCRQATLTLTQGP 180
DB 121 VILHTDSRKDSPPGSPAGRSINYSFYVYCKGPCQORVOPGKLRVQGSTCRQATLTLTQGP 180
QY 181 SCMDVLIPIRMSEGCSPHCPGTSAEFFPKCGAHPSTDKETPVLAHLIATNSRNTTCT 240
DB 181 SCMDVLIPIRMSEGCSPHCPGTSAEFFPKCGAHPSTDKETPVLAHLIATNSRNTTCT 240
QY 241 CTDVRSVPLVFOQNSRHVITCLDCFHLVCTVRLNDROFVHDPOLGYSLPCVAGCPNSLKE 300
DB 241 CTDVRSVPLVFOQNSRHVITCLDCFHLVCTVRLNDROFVHDPOLGYSLPCVAGCPNSLKE 300
QY 301 LHHFRILGEEQYNNRYOQYGAEECVLQMGVLCPRPGCGAGLLPEPDQRVTCGCGNGLGC 360
DB 301 LHHFRILGEEQYNNRYOQYGAEECVLQMGVLCPRPGCGAGLLPEPDQRVTCGCGNGLGC 360
QY 361 GFAPCRECKEAYHEGECSAVFEASGTTTQAYRVDEBAQAQAEASKEITIKTTKPCPR 420
DB 361 GFAPCRECKEAYHEGECSAVFEASGTTTQAYRVDEBAQAQAEASKEITIKTTKPCPR 420
QY 421 CHVPVEKNGGCMHKCPQPCRLIEWCMNCGCEWNRVCMGDHWFVDV 465
DB 421 CHVPVEKNGGCMHKCPQPCRLIEWCMNCGCEWNRVCMGDHWFVDV 465
```

RESULT 5
ID AAY32502 standard; protein; 437 AA.
XX AAY32502;

DT 21-OCT-1999 (first entry)

DE Human parkin gene variant protein.

KM Parkinson's disease related gene; parkin gene; variant; gene therapy.

XX Homo sapiens.

XX WO9940191-A1.

XX 12-AUG-1999.

XX 09-FEB-1999; 99WO-JP000545.

XX 09-FEB-1998; 98JP-00027531.

XX (SHIMU) SHIMIZU N.

PA (MIZU/) MIZUNO Y.

XX Shimizu N, Mizuno Y;

XX WPI; 1999-494295/41.

DR N-PSDB; AAX99224.

PT Gene implicated in the pathology of Parkinson's disease, used for
FT treatment of the disease.

XX Claim 1; Page 89-94; 11app; English.

CC This sequence is encoded by a gene of the invention, and is implicated in
CC the pathology of Parkinson's disease. This sequence is a variant of the
CC parkin gene found in parkinson's disease patients. The sequences may be
CC used for the diagnosis, treatment (including gene therapy) and
CC investigation of Parkinson's disease

SQ Sequence 437 AA;

Query Match 92.6%; Score 2405; DB 2; Length 437;
Best Local Similarity 94.0%; Pred. No. 1.3e-202;

Matches 437; Conservative 0; Mismatches 28; Indels 28; Gaps 1;

```
QY 1 MIVFRRFNSHGFPVEVSDTISIFOLKEVAKRGQVADQRLVIFAGKELRNDMTVQNC 60
DB 1 MIVFRRFNSHGFPVEVSDTISIFOLKEVAKRGQVADQRLVIFAGKELRNDMTVQNC 60
QY 61 LDQOSIVHIVORPMRKGEENATGDDPRNAAAGCEREPOSITRVLDSSSVLPEDSVGLA 120
DB 61 LDQOSIVHIVORPMRKGEENATGDDPRNAAAGCEREPOSITRVLDSSSVLPEDSVGLA 120
QY 121 VILHTDSRKDSPPGSPAGRSINYSFYVYCKGPCQORVOPGKLRVQGSTCRQATLTLTQGP 180
DB 121 VILHTDSRKDSPPGSPAGRSINYSFYVYCKGPCQORVOPGKLRVQGSTCRQATLTLTQGP 180
QY 181 SCMDVLIPIRMSEGCSPHCPGTSAEFFPKCGAHPSTDKETPVLAHLIATNSRNTTCT 240
DB 179 -----EFPKCGAHPSTDKETPVLAHLIATNSRNTTCT 212
QY 241 CTDVRSVPLVFOQNSRHVITCLDCFHLVCTVRLNDROFVHDPOLGYSLPCVAGCPNSLKE 300
DB 213 CTDVRSVPLVFOQNSRHVITCLDCFHLVCTVRLNDROFVHDPOLGYSLPCVAGCPNSLKE 272
QY 301 LHHFRILGEEQYNNRYOQYGAEECVLQMGVLCPRPGCGAGLLPEPDQRVTCGCGNGLGC 360
DB 273 LHHFRILGEEQYNNRYOQYGAEECVLQMGVLCPRPGCGAGLLPEPDQRVTCGCGNGLGC 332
QY 361 GFAPCRECKEAYHEGECSAVFEASGTTTQAYRVDEBAQAQAEASKEITIKTTKPCPR 420
DB 333 GFAPCRECKEAYHEGECSAVFEASGTTTQAYRVDEBAQAQAEASKEITIKTTKPCPR 392
QY 421 CHVPVEKNGGCMHKCPQPCRLIEWCMNCGCEWNRVCMGDHWFVDV 465
DB 393 CHVPVEKNGGCMHKCPQPCRLIEWCMNCGCEWNRVCMGDHWFVDV 437
```

RESULT 6
ID ABO07157 standard; protein; 437 AA.
XX ABO07157;

DT 13-AUG-2003 (first entry)

DE Human p53 modifying protein, SEQ ID 117.

KM Human; p53 modifier; cytosolic; cancer; cytosolic; antiangiogenic;
KW antiapoptotic; p53 pathway; breast cancer; colon cancer; kidney cancer;
KW lung cancer; ovarian cancer; angiogenesis; cell cycle;
KW apoptotic disorder; cell proliferation disorder.

XX Homo sapiens.

XX MO200239122-A1.
PN
XX 12-DEC-2002.
PD
XX 03-JUN-2002; 2002MO-US017382.
PF
XX 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 15-FEB-2002; 2002US-0357253P.
XX (EXEL-) EXELIXIS INC.
PA
PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX
XX MPI: 2003-156859/15.
DR N-PSDB; ACD13333.
XX
XX Identifying modulators of the p53 pathway for use in treating apoptotic
PT or cell proliferation disorders, comprises screening for agents that
PT modulate activity of a human ortholog of genes that modify the p53
PT pathway in Drosophila.

Example 2; Page 401-402; 678bp; English.

CC The invention relates to identifying (M1) a candidate p53 pathway
CC modulating agent, by contacting an assay system comprising a purified HM
CC polypeptide (human orthologue of genes that modify the p53 pathway in
CC Drosophila) or nucleic acid with a test agent under conditions, where but
CC for the presence of the test agent, the system provides a reference
CC activity, and detecting a test agent-biased activity of the assay system.
CC Also included are modulating (M2) a p53 pathway of a cell (comprising
CC contacting a cell defective in p53 function with a candidate modulator
CC that specifically binds to a HM polypeptide comprising an HM amino acid
CC sequence, where p53 function is restored), modulating (M3) a p53 pathway
CC in a mammalian cell (comprising contacting the cell with an agent that
CC specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
CC a disease in a patient (comprising: (a) obtaining a biological sample
CC from the patient; (b) contacting the sample with a probe for HM
CC expression; (c) comparing the results with a control; and (d) determining
CC whether the comparison indicates a likelihood disease). (M1) is useful
CC for identifying modulators of the p53 pathway. A probe for HM expression
CC is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,
CC in a patient, where the cancer has greater than 25 % expression level.
CC Modulators identified by (M1) are useful in a variety of diagnostic and
CC therapeutic applications, where disease or disorder prognosis is related
CC to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
CC proliferation disorders (e.g. cancer). Another two methods (M2 and
CC M3) are useful for modulating the p53 pathway of a cell, thus restoring
CC the p53 function of the cell, so that the cell undergoes normal
CC proliferation or progression through the cell cycle. (M2) and (M3) are
CC also useful for treating defects in the p53 pathway such as angiogenic,
CC apoptotic or cell proliferation disorders. The present sequence
CC represents a human p53 pathway modifying protein
XX
XX

Sequence 437 AA;

Query Match 92.6%; Score 2405; DB 6; Length 437;

Best Local Similarity 94.0%; Pred. No. 1.3e-202;

Matches 437; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

QY 1 MIVPFRFNSHGPPVSDSTSFOLKEVAVAKQVPADOLRIYFAGKELRNDMTQONDD 60
DB 1 MIVPFRFNSHGPPVSDSTSFOLKEVAVAKQVPADOLRIYFAGKELRNDMTQONDD 60
QY 61 LDOQSIIVHIVORPWKQCEMNATGDDPRNAGGCGEREPQSLTRVLDSSVLPJGDSVGLA 120
DB 61 LDOQSIIVHIVORPWKQCEMNATGDDPRNAGGCGEREPQSLTRVLDSSVLPJGDSVGLA 120
QY 121 VILHTDSRSDSPAGSPAGRSIYNSFYVCKGFCQVQVQPKLRVQGSTGRQATLTITGCP 180
DB 121 VILHTDSRSDSPAGSPAGRSIYNSFYVCKGFCQVQVQPKLRVQGSTGRQATLTITGCP 180
QY 121 VILHTDSRSDSPAGSPAGRSIYNSFYVCKGFCQVQVQPKLRVQGSTGRQATLTITGCP 180
DB 121 VILHTDSRSDSPAGSPAGRSIYNSFYVCKGFCQVQVQPKLRVQGSTGRQATLTITGCP 180

QY 181 SCMDVLLIPNRMSGECQSPHCPTGSAEFFPKCGAKPTSKETPVALLIATNSRNTCTT 240
DB 179 -----EFFPKCGAKPTSKETPVALLIATNSRNTCTT 212
QY 241 CTDVRSPTLVFOCNSRHVCTLDCPFHLYCYTRINDQFVNDPQGLVSLPCVAGCPNLLIE 300
DB 213 CTDVRSPTLVFOCNSRHVCTLDCPFHLYCYTRINDQFVNDPQGLVSLPCVAGCPNLLIE 272
QY 301 LHHFRILGEEQYNNRYQQYGAEECVLQMGVLCPRPCGAGLLPEPQRYVTCEGGNGLGC 360
DB 273 LHHFRILGEEQYNNRYQQYGAEECVLQMGVLCPRPCGAGLLPEPQRYVTCEGGNGLGC 332
QY 361 GFAPCRCKEAYHGECSAVFEASGTTQAYVNDERBAEQARWEAASKETIKKTKPCPR 420
DB 333 GFAPCRCKEAYHGECSAVFEASGTTQAYVNDERBAEQARWEAASKETIKKTKPCPR 392
QY 421 CHVPVEKXGSCMHMKCPQPCRLLEWCMNCGCEMNRVCMGDHWPV 465
DB 393 CHVPVEKXGSCMHMKCPQPCRLLEWCMNCGCEMNRVCMGDHWPV 437

RESULT 7

AAB67517
ID AAB67517 standard; protein; 464 AA.

AC AAB67517;

DT 29-MAY-2001 (first entry)

DE Amino acid sequence of a murine parkin2 polypeptide.

XX Parkinson's disease; 6q25.2-27; neurodegenerative disease;

KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;

KW Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;

KW Brain tumour; head trauma; stroke; vascular irregularity;

XX metabolic irregularity.

XX

OS Mus sp.

XX

PN EP1081225-A1.

XX

PD 07-MAR-2001.

XX

PF 30-AUG-1999; 99EP-00116766.

XX

PR 30-AUG-1999; 99EP-00116766.

XX

PA (BIOF-) BIOFRONTIERA PHARM GMBH.

XX

PI Luebbert H;

XX

PT MPI; 2001-212797/22.

DR N-PSDB; AAF55244.

XX

PT New polynucleotides encoding mouse parkin2 protein, useful for producing

PT a transgenic non-human animal as an animal model for neurodegenerative

PT diseases.

PS Disclosure; Page 17-19; 62pp; English.

XX

XX The present sequence represents a murine parkin2 polypeptide. Mutations

CC or deletions in the parkin2 gene cause Parkinson's disease in humans. The

CC human parkin2 gene is located in gene region 6q25.2-27. Parkin2

CC polypeptides and polynucleotides are useful for analysing

CC neurodegenerative diseases. They are also useful for testing the efficacy

CC of the treatment of a neurodegenerative disease such as Parkinson's

CC disease, Alzheimer's disease, Huntington's disease, amyotrophic lateral

CC sclerosis, Multi-system atrophy, Wilson's disease, Pick's disease, Prion

CC disease, and secondary causes inducing Parkinson's syndromes like toxins,

CC drugs, brain tumours, head trauma, stroke, vascular irregularities or

CC metabolic irregularities, associated with a less active or non-active

CC parkin protein

XX

SQ Sequence 464 AA;
 Query Match 83.6%; Score 2170.5; DB 4; Length 464;
 Best Local Similarity 83.4%; Pred. No. 66-182;
 Matches 388; Conservative 29; Mismatches 47; Indels 1; Gaps 1;

QY 1 MIVFVRNNSHGPFPEVDSSTSI FOLKEVAKRGQVPADQLRVIFAGKELRNDWTQVNC D 60
 DB 1 MIVFVRNNSHGPFPEVDSSTSI FOLKEVAKRGQVPADQLRVIFAGKELRNDWTQVNC D 60
 QY 61 LDQGSIVHIVORPWRKGEENATGDDPRNNAAGCEREPOSITRVDSSTVLPDSDVGLA 120
 DB 61 LDQGSIVHIVORPWRKGEENATGDDPRNNAAGCEREPOSITRVDSSTVLPDSDVGLA 120
 QY 121 VILHTDSKRDSPRAGSPAGRSIYNSFYVYCKGRCQVOPGLRVQCSCTCQATLTLLTQGP 180
 DB 121 VILHTDSKRDSEARGPV-KPTVNSFFIYCKGRCQVOPGLRVQCSCTCQATLTLLTQGP 179
 QY 181 SCMDVLLPNRMSEGCSPHCPGTSAEFFPKCGAHPISDKETPVALLIATNSRNTCTT 240
 DB 180 SCMDVLLPNRMSEGCSPHCPGTSAEFFPKCGAHPISDKETPVALLIATNSRNTCTT 239
 QY 241 CTDVRSPLVFOQNSRHVITCLDFHLYCVTRLNDROFVHDPOGLGYSLPCVAGCPSNLIKE 300
 DB 240 CTDVRSPLVFOQHHRVITCLDFHLYCVTRLNDROFVHDPOGLGYSLPCVAGCPSNLIKE 299
 QY 301 LHHFRILGEOYRNYOQYGAEECVLQMGVLCPRPGCGAGLPEPDQRKVTCEGGNGLGC 360
 DB 300 LHHFRILGEOYTRYOQYGAEECVLQMGVLCPRPGCGAGLPEPDQRKVTCEGGNGLGC 359
 QY 361 GFVFCRCKEAYHGECSAVFEASGTTQAYRVDERAAEQARWEASKEITIKTTTPCPR 420
 DB 360 GFVFCRCKEAYHGECDLSLEPSGATSOAYRVDKRAEQARWEASKEITIKTTTPCPR 419
 QY 421 CHVPEKNGGCMAMKCPQPCRLKWCWNGCGCEMNRYCMGDHMFV 465
 DB 420 CNVPIEKNGGCMAMKCPQPCRLKWCWNGCGCEMNRYCMGDHMFV 464

RESULT 8
 AAB67531 ID AAB67531 standard; protein; 464 AA.
 AC AAB67531;
 DT 29-MAY-2001 (first entry)
 DE Amino acid sequence of a mutated murine parkin2 polypeptide.
 KW Parkinson's disease; 6q25.2-27; neurodegenerative disease;
 KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
 KW Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
 KW brain tumour; head trauma; stroke; vascular irregularity;
 KW metabolic irregularity.
 OS Mus sp.
 XX
 XX
 PN EP1081225-A1.
 XX
 PD 07-MAR-2001.
 XX
 PF 30-AUG-1999; 99EP-00116766.
 XX
 PR 30-AUG-1999; 99EP-00116766.
 XX
 PA (BIOF-) BIOFRONTIERA PHARM GMBH.
 XX
 PI Luebbert H;
 XX
 XX WPI; 2001-212797/22.
 DR N-PsDB; AAF55258.
 XX
 PT New polynucleotides encoding mouse parkin2 protein, useful for producing

PT a transgenic non-human animal as an animal model for neurodegenerative
 PT diseases.
 PS
 SS Claim 7; Page 47-49; 62pp; English.
 XX
 CC The present sequence represents a murine parkin2 polypeptide. The
 CC sequence contains the mutation Lys161Asn. Mutations or deletions in the
 CC parkin2 gene cause Parkinson's disease in humans. The human parkin2 gene
 CC is located in gene region 6q25.2-27. Parkin2 polypeptides and
 CC polynucleotides are useful for analysing neurodegenerative diseases. They
 CC are also useful for testing the efficacy of the treatment of a
 CC neurodegenerative disease such as Parkinson's disease, Alzheimer's
 CC disease, Huntington's disease, amyotrophic lateral sclerosis, Multi-
 CC system atrophy, Wilson's disease, Pick's disease, Prion disease, and
 CC secondary causes inducing Parkinson's syndromes like toxins, drugs, brain
 CC tumours, head trauma, stroke, vascular irregularities or metabolic
 CC irregularities, associated with a less active or non-active parkin
 CC protein
 CC
 XX

SQ Sequence 464 AA;
 Query Match 83.4%; Score 2165.5; DB 4; Length 464;
 Best Local Similarity 83.2%; Pred. No. 1.7e-181;
 Matches 387; Conservative 29; Mismatches 48; Indels 1; Gaps 1;

QY 1 MIVFVRNNSHGPFPEVDSSTSI FOLKEVAKRGQVPADQLRVIFAGKELRNDWTQVNC D 60
 DB 1 MIVFVRNNSHGPFPEVDSSTSI FOLKEVAKRGQVPADQLRVIFAGKELRNDWTQVNC D 60
 QY 61 LDQGSIVHIVORPWRKGEENATGDDPRNNAAGCEREPOSITRVDSSTVLPDSDVGLA 120
 DB 61 LDQGSIVHIVORPWRKGEENATGDDPRNNAAGCEREPOSITRVDSSTVLPDSDVGLA 120
 QY 121 VILHTDSKRDSPRAGSPAGRSIYNSFYVYCKGRCQVOPGLRVQCSCTCQATLTLLTQGP 180
 DB 121 VILHTDSKRDSEARGPV-KPTVNSFFIYCKGRCQVOPGLRVQCSCTCQATLTLLTQGP 179
 QY 181 SCMDVLLPNRMSEGCSPHCPGTSAEFFPKCGAHPISDKETPVALLIATNSRNTCTT 240
 DB 180 SCMDVLLPNRMSEGCSPHCPGTSAEFFPKCGAHPISDKETPVALLIATNSRNTCTT 239
 QY 241 CTDVRSPLVFOQNSRHVITCLDFHLYCVTRLNDROFVHDPOGLGYSLPCVAGCPSNLIKE 300
 DB 240 CTDVRSPLVFOQHHRVITCLDFHLYCVTRLNDROFVHDPOGLGYSLPCVAGCPSNLIKE 299
 QY 301 LHHFRILGEOYRNYOQYGAEECVLQMGVLCPRPGCGAGLPEPDQRKVTCEGGNGLGC 360
 DB 300 LHHFRILGEOYTRYOQYGAEECVLQMGVLCPRPGCGAGLPEPDQRKVTCEGGNGLGC 359
 QY 361 GFVFCRCKEAYHGECSAVFEASGTTQAYRVDERAAEQARWEASKEITIKTTTPCPR 420
 DB 360 GFVFCRCKEAYHGECDLSLEPSGATSOAYRVDKRAEQARWEASKEITIKTTTPCPR 419
 QY 421 CHVPEKNGGCMAMKCPQPCRLKWCWNGCGCEMNRYCMGDHMFV 465
 DB 420 CNVPIEKNGGCMAMKCPQPCRLKWCWNGCGCEMNRYCMGDHMFV 464

RESULT 9
 AAB67532 ID AAB67532 standard; protein; 464 AA.
 AC AAB67532;
 DT 29-MAY-2001 (first entry)
 DE Amino acid sequence of a mutated murine parkin2 polypeptide.
 KW Parkinson's disease; 6q25.2-27; neurodegenerative disease;
 KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
 KW Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
 KW brain tumour; head trauma; stroke; vascular irregularity;
 KW metabolic irregularity.

Db 121 VILDTDSKRDSBARGPV-KPTVNSFFIYCKGPHKVGPKLKVQCGTCKQATLTLAQP 179
 Qy 181 SCMDVLIIPRMSSGECSPHCPTSAEFPFKCAHPTSDKTEPVALLHLATNSRNTTCT 240
 Db 180 SCMDVLIIPRMSSGECSPHCPTSAEFPFKCAHPTSDKTSVAMNTITSRRSIPCTA 239
 Qy 241 CTDVRSPLVIFQCNRSRNVICLDCPHLYCVTRILNDROFVHDPOLGYSLPCVAGCPNSLKE 300
 Db 240 CTDVRSPLVIFQCNRSRNVICLDCPHLYCVTRILNDROFVHDPOLGYSLPCVAGCPNSLKE 299
 Qy 301 LHHFRILIGEEQYRNYOQYGAEECVLQMGVLCPRPGCAGLLPEPDQRKVTCEGNGIGC 360
 Db 300 LHHFRILIGEEQYRNYOQYGAEECVLQMGVLCPRPGCAGLLPEPDQRKVTCEGNGIGC 359
 Qy 361 GFAPCRECKEAYHGECSAVFEASGTTTQAYRVDERAAEOARWEASKEITIKTTKPCPR 420
 Db 360 GFAPCRECKEAYHGECSAVFEASGTTTQAYRVDERAAEOARWEASKEITIKTTKPCPR 419
 Qy 421 CHVPEKNGGCMHMKCPQPCRLKWCNMNGCE 452
 Db 420 CNVPEKNGGCMHMKCPQPCRLKWCNMNGCE 451

RESULT 11

ABO07158

ID ABO07158 standard; protein; 316 AA.

XX AC ABO07158;

DT 13-AUG-2003 (first entry)

DE Human p53 modifying protein, SEQ ID 118.

KM Human; p53 modifier; cytostatic; cancer; cytostatic; antiangiogenic;
 KM antiapoptotic; p53 pathway; breast cancer; colon cancer; kidney cancer;
 KM lung cancer; ovarian cancer; angiogenesis; cell cycle;
 KM apoptotic disorder; cell proliferation disorder.

OS Homo sapiens.

XX XX W0200299122-A1.

XX PD 12-DEC-2002.

XX PF 03-JUN-2002; 2002M0-US017382.

XX PR 05-JUN-2001; 2001US-0296076P.

XX PR 10-OCT-2001; 2001US-0328605P.

XX PR 15-FEB-2002; 2002US-0357253P.

XX PA (EXEL-) EXELIXIS INC.

XX PI Friedman L, Florman GD, Belvin M, Francis-Lang H, Li D, Funke RP;

XX DR WPI, 2003-156859/15.

XX DR N-PSDB; ACD13334.

PT Identifying modulators of the p53 pathway for use in treating apoptotic
 PT or cell proliferation disorders, comprises screening for agents that
 PT modulate activity of a human ortholog of genes that modify the p53
 PT pathway in Drosophila.

XX XX Example 2; Page 402-403; 678bp; English.

CC The invention relates to identifying (M1) a candidate p53 pathway
 CC modulating agent, by contacting an assay system comprising a purified HM
 CC polypeptide (human orthologue of genes that modify the p53 pathway in
 CC Drosophila) or nucleic acid with a test agent under conditions, where but
 CC for the presence of the test agent, the system provides a reference
 CC activity, and detecting a test agent-biased activity of the assay system.
 CC Also included are modulating (M2) a p53 pathway of a cell comprising
 CC contacting a cell defective in p53 function with a candidate modulator
 CC that specifically binds to a HM polypeptide comprising an HM amino acid

CC sequence, where p53 function is restored), modulating (M3) a p53 pathway
 CC in a mammalian cell (comprising contacting the cell with an agent that
 CC specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
 CC a disease in a patient (comprising: (a) obtaining a biological sample
 CC from the patient; (b) contacting the sample with a probe for HM
 CC expression; (c) comparing the results with a control; and (d) determining
 CC whether the comparison indicates a likelihood disease). (M1) is useful
 CC for identifying modulators of the p53 pathway. A probe for HM expression
 CC is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,
 CC in a patient, where the cancer has greater than 25 % expression level.
 CC Modulators identified by (M1) are useful in a variety of diagnostic and
 CC therapeutic applications, where disease or disorder prognosis is related
 CC to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
 CC proliferation disorders (e.g. cancer). Another two new methods (M2 and
 CC M3) are useful for modulating the p53 pathway of a cell, thus restoring
 CC the p53 function of the cell, so that the cell undergoes normal
 CC proliferation or progression through the cell cycle. (M2) and (M3) are
 CC also useful for treating defects in the p53 pathway such as angiogenic,
 CC apoptotic or cell proliferation disorders. The present sequence
 CC represents a human p53 pathway modifying protein

XX S0 Sequence 316 AA;

Query Match 65.4%; Score 1696.5; DB 6; Length 316;

Best Local Similarity 68.0%; Pred. No. 1.8e-140; Mismatches 1;

Matches 316; Conservative 0; Mismatches 0; Indels 149; Gaps 1;

Qy 1 MIVFVRENSHGFPVVDSDTSIFOLKEVNAKROGVPADQLRVIFAGKEIRNDVTYONCD 60

Db 1 MIVFVRENSHGFPVVDSDTSIFOLKEVNAKROGVPADQLRVIFAGKEIRNDVTYONCD 57

Qy 61 LDQGSIVHIVPRKQEMNATGDDPRNAAGCEREPQSLRTVLDLSSVLPGDSVGLA 120

Db 58 ----- 57

Qy 121 VILHTRKDSPPAGSPAGRSIYNSFYVYCKPGQVRVQPKLRVQGSTGRQATLTLTQPR 180

Db 58 ----- 57

Qy 181 SCMDVLIIPRMSSGECSPHCPTSAEFPFKCAHPTSDKTEPVALLHLATNSRNTTCT 240

Db 58 -----EFPFKCAHPTSDKTEPVALLHLATNSRNTTCT 91

Qy 241 CTDVRSPLVIFQCNRSRNVICLDCPHLYCVTRILNDROFVHDPOLGYSLPCVAGCPNSLKE 300

Db 92 CTDVRSPLVIFQCNRSRNVICLDCPHLYCVTRILNDROFVHDPOLGYSLPCVAGCPNSLKE 151

Qy 301 LHHFRILIGEEQYRNYOQYGAEECVLQMGVLCPRPGCAGLLPEPDQRKVTCEGNGIGC 360

Db 152 LHHFRILIGEEQYRNYOQYGAEECVLQMGVLCPRPGCAGLLPEPDQRKVTCEGNGIGC 211

Qy 361 GFAPCRECKEAYHGECSAVFEASGTTTQAYRVDERAAEOARWEASKEITIKTTKPCPR 420

Db 212 GFAPCRECKEAYHGECSAVFEASGTTTQAYRVDERAAEOARWEASKEITIKTTKPCPR 271

Qy 421 CHVPEKNGGCMHMKCPQPCRLKWCNMNGCEMNRCVMDHMFV 465

Db 272 CHVPEKNGGCMHMKCPQPCRLKWCNMNGCEMNRCVMDHMFV 316

RESULT 12

AAB67521

ID AAB67521 standard; protein; 344 AA.

XX AC AAB67521;

DT 29-MAY-2001 (first entry)

DE Amino acid sequence of a murine truncated parkin2 polypeptide.

KM Parkinson's disease; 6q25.2-27; neurodegenerative disease;
 KM Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
 KM Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;

KW brain tumour; head trauma; stroke; vascular irregularity;
 KW metabolic irregularity.
 OS Mus sp.
 XX EPI081225-A1.
 XX
 XX PD 07-MAR-2001.
 XX
 XX PF 30-AUG-1999; 99EP-00116766.
 XX
 XX PR 30-AUG-1999; 99EP-00116766.
 XX
 XX PA (BIOF-) BIOFRONTERA PHARM GMBH.
 XX
 XX PI Luebert H;
 XX
 XX DR WPI; 2001-212797/22.
 XX DR N-PSDB; AAF55248.
 XX
 PT New polynucleotides encoding mouse parkin2 protein, useful for producing
 PT a transgenic non-human animal as an animal model for neurodegenerative
 PT diseases.
 PS
 XX
 XX PS Claim 7; Page 40-41; 62pp; English.
 XX
 CC The present sequence represents a murine parkin2 polypeptide. The
 CC polynucleotide sequence contains a deletion, leading to a truncated
 CC protein. Mutations or deletions in the parkin2 gene cause Parkinson's
 CC disease in humans. The human parkin2 gene is located in gene region
 CC 6q25.2-27. Parkinson polypeptides and polynucleotides are useful for
 CC analysing neurodegenerative diseases. They are also useful for testing
 CC the efficacy of the treatment of a neurodegenerative disease such as
 CC Parkinson's disease, Alzheimer's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease,
 CC Pick's disease, Prion disease, and secondary causes inducing Parkinson's
 CC syndromes like toxins, drugs, brain tumours, head trauma, stroke,
 CC vascular irregularities or metabolic irregularities, associated with a
 CC less active or non-active parkin protein
 XX
 XX Sequence 344 AA;
 SQ
 Query Match 64.7%; Score 1679.5; DB 4; Length 344;
 Best Local Similarity 66.0%; Pred. No. 6.3e-139;
 Matches 307; Conservative 14; Mismatches 23; Indels 121; Gaps 1;

QY 1 MIVFVRENSHGFPVEVSDTSIFOLKEVAVARQGVADQLRVIFAGKELRNDWTYQNC 60
 DB 1 MIVFVRENSHGFPVEVSDTSIFOLKEVAVARQGVADQLRVIFAGKELRNDWTYQNC 60
 QY 61 LDOQSVIHTVQRPVRKQGMNATGGDDPRNNAAGCEREPQSLTRVDLSSVLRGDSVGLA 120
 DB 57 ----- 56
 QY 121 VIHTDSRKDSPAGSPAGRSIYNFVYCKGPQGVQPKLKVQSCSTGRQATLTLTGCP 180
 DB 57 ----- 56
 QY 57 ----- 59
 DB 181 SCWDVLIIPNRMSGSCQSPHCPGTSAPFFKCGAMPTSDKETPVALLIATNSRNTTCT 240
 DB 60 SCWDVLIIPNRMSGSCQSPHCPGTSAPFFKCGAMPTSDKETPVALLIATNSRNTTCT 240
 QY 241 CTDVNSPVVIFQCNHNVICLDCPHLYCYTRLNDRPFVNDPQLGYSIPCVAGCPNLLIE 300
 DB 120 CTDVNSPVVIFQCNHNVICLDCPHLYCYTRLNDRPFVNDPQLGYSIPCVAGCPNLLIE 300
 QY 301 LHHFRILGEEQYNNRYYQYGAEECVLQMGVLCRPGCGAGLLPEPDQRKYTCGGNGLGC 360
 DB 180 LHHFRILGEEQYNNRYYQYGAEECVLQMGVLCRPGCGAGLLPEPDQRKYTCGGNGLGC 360
 QY 361 GAFCPECKEAYHGECSAVFEASGTTTQAYRVDERBAABOARWEAASKETIKTKTTCPR 420
 DB 240 GAFCPECKEAYHGECSAVFEASGTTTQAYRVDERBAABOARWEAASKETIKTKTTCPR 420

QY 421 CHVPEKNGSCMHMKCPQPCRLIEWCMNCGCEWNRVCMGDHMDV 465
 DB 300 CHVPEKNGSCMHMKCPQPCRLIEWCMNCGCEWNRVCMGDHMDV 344
 RESULT 13
 AAB67526
 ID AAB67526 standard; protein; 296 AA.
 XX
 XX AC AAB67526;
 XX
 XX DT 29-MAY-2001 (first entry)
 XX
 XX DE Amino acid sequence of a murine truncated parkin2 polypeptide.
 XX
 KW Parkinson's disease; 6q25.2-27; neurodegenerative disease;
 KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
 KW Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
 KW brain tumour; head trauma; stroke; vascular irregularity;
 KW metabolic irregularity.
 XX
 XX OS Mus sp.
 XX
 XX PN EPI081225-A1.
 XX
 XX PD 07-MAR-2001.
 XX
 XX PF 30-AUG-1999; 99EP-00116766.
 XX
 XX PR 30-AUG-1999; 99EP-00116766.
 XX
 XX PA (BIOF-) BIOFRONTERA PHARM GMBH.
 XX
 XX PI Luebert H;
 XX
 XX DR WPI; 2001-212797/22.
 XX DR N-PSDB; AAF55253.
 XX
 PT New polynucleotides encoding mouse parkin2 protein, useful for producing
 PT a transgenic non-human animal as an animal model for neurodegenerative
 PT diseases.
 PS
 XX
 XX PS Claim 7; Page 44-45; 62pp; English.
 XX
 CC The present sequence represents a murine parkin2 polypeptide. The
 CC polynucleotide sequence contains a deletion, leading to a truncated
 CC protein. Mutations or deletions in the parkin2 gene cause Parkinson's
 CC disease in humans. The human parkin2 gene is located in gene region
 CC 6q25.2-27. Parkinson polypeptides and polynucleotides are useful for
 CC analysing neurodegenerative diseases. They are also useful for testing
 CC the efficacy of the treatment of a neurodegenerative disease such as
 CC Parkinson's disease, Alzheimer's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease,
 CC Pick's disease, Prion disease, and secondary causes inducing Parkinson's
 CC syndromes like toxins, drugs, brain tumours, head trauma, stroke,
 CC vascular irregularities or metabolic irregularities, associated with a
 CC less active or non-active parkin protein
 XX
 XX Sequence 296 AA;
 SQ
 Query Match 47.6%; Score 1234.5; DB 4; Length 296;
 Best Local Similarity 80.0%; Pred. No. 7.3e-100;
 Matches 232; Conservative 20; Mismatches 37; Indels 1; Gaps 1;

QY 1 MIVFVRENSHGFPVEVSDTSIFOLKEVAVARQGVADQLRVIFAGKELRNDWTYQNC 60
 DB 1 MIVFVRENSHGFPVEVSDTSIFOLKEVAVARQGVADQLRVIFAGKELRNDWTYQNC 60
 QY 61 LDOQSVIHTVQRPVRKQGMNATGGDDPRNNAAGCEREPQSLTRVDLSSVLRGDSVGLA 120
 DB 61 LDOQSVIHTVQRPVRKQGMNATGGDDPRNNAAGCEREPQSLTRVDLSSVLRGDSVGLA 120

CC sclerosis, Multi-system atrophy, Wilson's disease, Prion
CC disease, and secondary causes inducing Parkinson's syndromes like toxins,
CC drugs, brain tumours, head trauma, stroke, vascular irregularities or
CC metabolic irregularities, associated with a less active or non-active
CC parkin protein

XX Sequence 250 AA;

Query Match 38.6%; Score 1002; DB 4; Length 250;

Best Local Similarity 77.6%; Pred. No. 1.7e-79;

Matches 190; Conservative 20; Mismatches 35; Indels 0; Gaps 0;

| | | | |
|----|-----|---|-----|
| Qy | 1 | MIVFVRFNSSHGFPVEVDSDTSIFQLKEVAVAKRGVPAQLRVIPAGKILRNDMTVONCD | 60 |
| | | : : : : : : : : | |
| Db | 1 | MIVFVRFNSSYGFPVEVDSDTSILQLKEVAVAKRGVPAQLRVIPAGKILPNHLTVQNC | 60 |
| | | : : : : : : : | |
| Qy | 61 | LDQGSIVHIVQRPWRKQGMNATGDDPRNNAAGCEREPOSITRYDLSISVLPDPSVGLA | 120 |
| | | : : : : : : : : | |
| Db | 61 | LEQGSIVHIVQRPWRSHETNAGSGDEPOSTSEGSIWESRSLTRYDLSISHTLPVDSVGLA | 120 |
| | | : : : : : : : : | |
| Qy | 121 | VILHTRSRKDSPPAGSPAGRSIYNSFYVYCKGPGCORVOPGKLRVOCSTCRQATLTLTQGP | 180 |
| | | : : : : : : : | |
| Db | 121 | VILDTSRKDSPPAGSPAGRSIYNSFYVYCKGPGCHKVQPGKLRVOCSTCRQATLTLTQGP | 180 |
| | | : : : : : : : | |
| Qy | 181 | SCMDVLLIPNRMSGECOSPHCPGTSAEFFKCAHPTSDKETPVALLIATNSRNITCIT | 240 |
| | | : : : : : : : | |
| Db | 181 | SCMDVLLIPNRMSGECOSPHCPGTSAEFFKCAHPTSDKETPVALLIATNSRNITCIT | 240 |
| | | : : : : : : : | |
| Qy | 241 | CTDVR 245 | |
| | | | |
| Db | 241 | CTDVR 245 | |

Search completed: February 23, 2005, 13:50:12
Job time : 105.258 secs

iris page blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:36:20 ; Search time 25.8801 seconds
(without alignments)
1341.256 Million cell updates/sec

Title: US-10-622-817-8

Perfect score: 2596
Sequence: 1 MIVFVRFNSSHGFPVEVDS.....CWNCGCEMNRVCMDHMDV 465

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|------------------------|--------------------|
| 1 | 2596 | 100.0 | 465 | 4 US-09-601-844B-2 | Sequence 2, Appl1 |
| 2 | 2596 | 100.0 | 465 | 4 US-09-949-016-6516 | Sequence 6516, Ap |
| 3 | 2405 | 92.6 | 437 | 4 US-09-601-844B-4 | Sequence 4, Appl1 |
| 4 | 2405 | 92.6 | 437 | 4 US-09-949-016-6853 | Sequence 6853, Ap |
| 5 | 1696.5 | 65.4 | 316 | 4 US-09-949-016-6854 | Sequence 6854, Ap |
| 6 | 370.5 | 14.3 | 117 | 4 US-09-270-767-32183 | Sequence 32183, A |
| 7 | 240.5 | 9.3 | 503 | 4 US-09-914-259-13 | Sequence 13, Appl1 |
| 8 | 222 | 8.6 | 474 | 4 US-09-354-221-2 | Sequence 2, Appl1 |
| 9 | 222 | 8.6 | 487 | 4 US-09-949-016-10111 | Sequence 10111, A |
| 10 | 191 | 7.4 | 506 | 4 US-09-248-796A-15410 | Sequence 15410, A |
| 11 | 167 | 6.4 | 587 | 1 US-08-398-008A-2 | Sequence 2, Appl1 |
| 12 | 167 | 6.4 | 587 | 2 US-08-893-333-2 | Sequence 2, Appl1 |
| 13 | 149 | 5.7 | 464 | 4 US-09-538-092-598 | Sequence 598, Appl |
| 14 | 141 | 5.4 | 328 | 4 US-09-248-796A-16471 | Sequence 16471, A |
| 15 | 133 | 5.1 | 352 | 4 US-08-854-764-2 | Sequence 2, Appl1 |
| 16 | 133 | 5.1 | 352 | 5 PCT-US95-09377-2 | Sequence 2, Appl1 |
| 17 | 131.5 | 5.1 | 1106 | 4 US-09-949-016-9626 | Sequence 9626, Ap |
| 18 | 130.5 | 5.0 | 1242 | 4 US-09-488-270A-2 | Sequence 2, Appl1 |
| 19 | 128.5 | 4.9 | 229 | 2 US-08-726-306A-23 | Sequence 23, Appl1 |
| 20 | 128.5 | 4.9 | 229 | 3 US-08-840-146-20 | Sequence 20, Appl1 |
| 21 | 128.5 | 4.9 | 229 | 3 US-09-360-220-20 | Sequence 20, Appl1 |
| 22 | 128 | 4.9 | 160 | 4 US-09-370-838-205 | Sequence 205, Appl |
| 23 | 128 | 4.9 | 160 | 4 US-09-854-133-205 | Sequence 205, Appl |
| 24 | 127 | 4.9 | 2321 | 4 US-09-230-652-2 | Sequence 2, Appl1 |
| 25 | 126 | 4.9 | 237 | 4 US-09-248-796A-15062 | Sequence 19062, A |
| 26 | 126 | 4.9 | 846 | 1 US-08-149-103-3 | Sequence 3, Appl1 |
| 27 | 126 | 4.9 | 846 | 1 US-08-451-883-3 | Sequence 3, Appl1 |

| | | | | | |
|----|-------|-----|------|------------------------|--------------------|
| 28 | 126 | 4.9 | 873 | 1 US-08-393-734-2 | Sequence 2, Appl1 |
| 29 | 126 | 4.9 | 904 | 3 US-08-894-489-2 | Sequence 2, Appl1 |
| 30 | 126 | 4.9 | 973 | 4 US-09-949-016-9528 | Sequence 9528, Ap |
| 31 | 125.5 | 4.8 | 103 | 2 US-08-771-201-9 | Sequence 9, Appl1 |
| 32 | 125 | 4.8 | 76 | 1 US-08-232-815-2 | Sequence 2, Appl1 |
| 33 | 125 | 4.8 | 76 | 1 US-08-350-906-2 | Sequence 2, Appl1 |
| 34 | 125 | 4.8 | 76 | 1 US-09-601-844B-61 | Sequence 61, Appl1 |
| 35 | 125 | 4.8 | 76 | 5 PCT-US95-04536-2 | Sequence 2, Appl1 |
| 36 | 125 | 4.8 | 206 | 4 US-09-248-796A-19430 | Sequence 19430, A |
| 37 | 125 | 4.8 | 1121 | 1 US-07-789-915A-2 | Sequence 2, Appl1 |
| 38 | 125 | 4.8 | 1121 | 1 US-08-005-002C-2 | Sequence 2, Appl1 |
| 39 | 125 | 4.8 | 1121 | 1 US-08-487-203A-2 | Sequence 2, Appl1 |
| 40 | 124.5 | 4.8 | 77 | 6 5510474-3 | Patent No. 5510474 |
| 41 | 124.5 | 4.8 | 77 | 6 5510474-3 | Patent No. 5510474 |
| 42 | 124.5 | 4.8 | 305 | 4 US-09-866-153-6 | Sequence 6, Appl1 |
| 43 | 124.5 | 4.8 | 305 | 4 US-09-693-467A-6 | Sequence 6, Appl1 |
| 44 | 124.5 | 4.8 | 305 | 4 US-09-270-976-6 | Sequence 6, Appl1 |
| 45 | 124.5 | 4.8 | 381 | 4 US-09-866-153-9 | Sequence 9, Appl1 |

ALIGNMENTS

| | | | | | |
|--|-----|--|--|-------------------------------------|-----|
| RESULT 1 | | | | | |
| US-09-601-844B-2 | | | | | |
| Sequence 2, Application US/09601844B | | | | | |
| Patent No. 6716621 | | | | | |
| GENERAL INFORMATION: | | | | | |
| APPLICANT: Shimizu, No. 6716621uyoshi | | | | | |
| APPLICANT: Mizuno, Yoshiaki | | | | | |
| TITLE OF INVENTION: Isolated DNA or Gene Responsible for Parkinson's Disease | | | | | |
| FILE REFERENCE: 0652-2110000 | | | | | |
| CURRENT APPLICATION NUMBER: US/09/601,844B | | | | | |
| PRIOR FILING DATE: 2000-08-09 | | | | | |
| PRIOR APPLICATION NUMBER: PCT/JP99/00545 | | | | | |
| PRIOR FILING DATE: 1999-02-09 | | | | | |
| NUMBER OF SEQ ID NOS: 70 | | | | | |
| SOFTWARE: Patentin version 3.1 | | | | | |
| SEQ ID NO 2 | | | | | |
| LENGTH: 465 | | | | | |
| TYPE: PRT | | | | | |
| ORGANISM: Homo sapiens | | | | | |
| US-09-601-844B-2 | | | | | |
| Query Match | | | | | |
| Best Local Similarity 100.0%; Score 2596; DB 4; Length 465; | | | | | |
| Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | | |
| QY | 1 | MIVFVRFNSSHGFPVEVDS | DTSI | FOLEKVVAKRGVPADQLRVTFAGKELRNDMTVQNC | 60 |
| DB | 1 | MIVFVRFNSSHGFPVEVDS | DTSI | FOLEKVVAKRGVPADQLRVTFAGKELRNDMTVQNC | 60 |
| QY | 61 | LDQOSTVHIVQRPWRKGGEMNATGDDPRNAAAGCEREPOS | LTRVDLSSSVLP | FGDSVGLA | 120 |
| DB | 61 | LDQOSTVHIVQRPWRKGGEMNATGDDPRNAAAGCEREPOS | LTRVDLSSSVLP | FGDSVGLA | 120 |
| QY | 121 | VILHTDSRSDSPAGSPAGRS | TVNSFVYVCKPCCQVQVQGXLRVQSTORQATLT | TOGP | 180 |
| DB | 121 | VILHTDSRSDSPAGSPAGRS | TVNSFVYVCKPCCQVQVQGXLRVQSTORQATLT | TOGP | 180 |
| QY | 121 | VILHTDSRSDSPAGSPAGRS | TVNSFVYVCKPCCQVQVQGXLRVQSTORQATLT | TOGP | 180 |
| DB | 121 | VILHTDSRSDSPAGSPAGRS | TVNSFVYVCKPCCQVQVQGXLRVQSTORQATLT | TOGP | 180 |
| QY | 181 | SCMDVLIENRSGEQSPHC | PGTSAEPFKGCAHPTSDKETPVVAHLAT | ATNSRNTTCT | 240 |
| DB | 181 | SCMDVLIENRSGEQSPHC | PGTSAEPFKGCAHPTSDKETPVVAHLAT | ATNSRNTTCT | 240 |
| QY | 241 | CTDVRSPLVLFQCN | SRHVICLDCPHLYCYTRINDRQFVADPOLGYSLPCVAGCPNS | LIKE | 300 |
| DB | 241 | CTDVRSPLVLFQCN | SRHVICLDCPHLYCYTRINDRQFVADPOLGYSLPCVAGCPNS | LIKE | 300 |
| QY | 301 | LHHFRILGEOVNR | VOYGAEECVLQMGVULCPRECGAGLLPEP | ORVYTCGGNGLGC | 360 |
| DB | 301 | LHHFRILGEOVNR | VOYGAEECVLQMGVULCPRECGAGLLPEP | ORVYTCGGNGLGC | 360 |
| QY | 361 | GFAFCREKCAVH | EGCSAVFEASGTTTQAYRVDERAALQAAWEAS | SKETIKKTKAPCR | 420 |
| DB | 361 | GFAFCREKCAVH | EGCSAVFEASGTTTQAYRVDERAALQAAWEAS | SKETIKKTKAPCR | 420 |

Db 361 GFAFCECKEAYHEGCSAVFEASGTTTAYVDERAAEQARWEASKEITTKTKPCR 420
QY 421 CHVPEKNGGCMHMKCPQPCRLKEMCNMGCEMNRVCMGDHMFV 465
Db 421 CHVPEKNGGCMHMKCPQPCRLKEMCNMGCEMNRVCMGDHMFV 465

RESULT 2

US-09-949-016-6516
Sequence 6516, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949, 016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 6516
LENGTH: 465
TYPE: PRT
ORGANISM: Human
US-09-949-016-6516

Query Match 100.0%; Score 2596; DB 4; Length 465;
Best Local Similarity 100.0%; Pred. No. 8.4e-251;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIVFVFNSSHGFPVEVDSDTSTIFOLKEVAKRGQVADQLRVIFAGKEIRNDMTVONCD 60
Db 1 MIVFVFNSSHGFPVEVDSDTSTIFOLKEVAKRGQVADQLRVIFAGKEIRNDMTVONCD 60
QY 61 LDQOSIVHIVQPRMKGEMNATGDDPRNAAAGCEKEPQSLTRVDSLSSVLPQDSVGLA 120
Db 61 LDQOSIVHIVQPRMKGEMNATGDDPRNAAAGCEKEPQSLTRVDSLSSVLPQDSVGLA 120
QY 121 VILHTDSKSDPPAGSPAGRSIYNSFYVCKGFCQGVQPKLRVQSTCRQATLTITQGP 180
Db 121 VILHTDSKSDPPAGSPAGRSIYNSFYVCKGFCQGVQPKLRVQSTCRQATLTITQGP 180
QY 181 SCMDVLIIPNMSGECSPHCPGTSAEFFPKCGAHTSDKETPVALLIATNSRNITCIT 240
Db 181 SCMDVLIIPNMSGECSPHCPGTSAEFFPKCGAHTSDKETPVALLIATNSRNITCIT 240
QY 241 CTDVRSFVLVFCQNSRHVICLDCFHLVCTRLNDRQFVADPOLGYSILPCVAGCPNSLIKE 300
Db 241 CTDVRSFVLVFCQNSRHVICLDCFHLVCTRLNDRQFVADPOLGYSILPCVAGCPNSLIKE 300
QY 301 LHHFRILGEBQYRKYQYGAEECVLQMGSVLCPRPGCGAGLLPEBDQKVTCEGNGIGC 360
Db 301 LHHFRILGEBQYRKYQYGAEECVLQMGSVLCPRPGCGAGLLPEBDQKVTCEGNGIGC 360
QY 361 GFAFCECKEAYHEGCSAVFEASGTTTQAYVDERAAEQARWEASKEITTKTKPCR 420
Db 361 GFAFCECKEAYHEGCSAVFEASGTTTQAYVDERAAEQARWEASKEITTKTKPCR 420
QY 421 CHVPEKNGGCMHMKCPQPCRLKEMCNMGCEMNRVCMGDHMFV 465
Db 421 CHVPEKNGGCMHMKCPQPCRLKEMCNMGCEMNRVCMGDHMFV 465

RESULT 3

US-09-601-844B-4
Sequence 4, Application US/09601844B

Patent No. 6716621
GENERAL INFORMATION:
APPLICANT: Shimizu, No. 6716621yoshi
APPLICANT: Mizuno, Yoshikuni
TITLE OF INVENTION: Isolated DNA or Gene Responsible for Parkinson's Disease
FILE REFERENCE: 0652, 211000
CURRENT APPLICATION NUMBER: US/09/601, 844B
PRIOR FILING DATE: 2000-08-09
PRIOR APPLICATION NUMBER: PCT/JP99/00545
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 437
TYPE: PRT
ORGANISM: Homo sapiens
US-09-601-844B-4

Query Match 92.6%; Score 2405; DB 4; Length 437;
Best Local Similarity 94.0%; Pred. No. 9.6e-232;
Matches 437; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

QY 1 MIVFVFNSSHGFPVEVDSDTSTIFOLKEVAKRGQVADQLRVIFAGKEIRNDMTVONCD 60
Db 1 MIVFVFNSSHGFPVEVDSDTSTIFOLKEVAKRGQVADQLRVIFAGKEIRNDMTVONCD 60
QY 61 LDQOSIVHIVQPRMKGEMNATGDDPRNAAAGCEKEPQSLTRVDSLSSVLPQDSVGLA 120
Db 61 LDQOSIVHIVQPRMKGEMNATGDDPRNAAAGCEKEPQSLTRVDSLSSVLPQDSVGLA 120
QY 121 VILHTDSKSDPPAGSPAGRSIYNSFYVCKGFCQGVQPKLRVQSTCRQATLTITQGP 180
Db 121 VILHTDSKSDPPAGSPAGRSIYNSFYVCKGFCQGVQPKLRVQSTCRQATLTITQGP 180
QY 181 SCMDVLIIPNMSGECSPHCPGTSAEFFPKCGAHTSDKETPVALLIATNSRNITCIT 240
Db 179 -----EFPFKCGAHTSDKETPVALLIATNSRNITCIT 212
QY 241 CTDVRSFVLVFCQNSRHVICLDCFHLVCTRLNDRQFVADPOLGYSILPCVAGCPNSLIKE 300
Db 213 CTDVRSFVLVFCQNSRHVICLDCFHLVCTRLNDRQFVADPOLGYSILPCVAGCPNSLIKE 272
QY 301 LHHFRILGEBQYRKYQYGAEECVLQMGSVLCPRPGCGAGLLPEBDQKVTCEGNGIGC 360
Db 273 LHHFRILGEBQYRKYQYGAEECVLQMGSVLCPRPGCGAGLLPEBDQKVTCEGNGIGC 332
QY 361 GFAFCECKEAYHEGCSAVFEASGTTTQAYVDERAAEQARWEASKEITTKTKPCR 420
Db 333 GFAFCECKEAYHEGCSAVFEASGTTTQAYVDERAAEQARWEASKEITTKTKPCR 392
QY 421 CHVPEKNGGCMHMKCPQPCRLKEMCNMGCEMNRVCMGDHMFV 465
Db 393 CHVPEKNGGCMHMKCPQPCRLKEMCNMGCEMNRVCMGDHMFV 437

RESULT 4

US-09-949-016-6853
Sequence 6853, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949, 016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6853
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6853

Query Match 92.6%; Score 2405; DB 4; Length 437;
Best Local Similarity 94.0%; Pred. No. 9, 6e-232;
Matches 437; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

```
QY 1 MIVFVRPNSSHGFPVEVDSITSIFOLKEVYAKRGVPADQLAVIFAGKELRNDWTQVQND 60
DB 1 MIVFVRPNSSHGFPVEVDSITSIFOLKEVYAKRGVPADQLAVIFAGKELRNDWTQVQND 60
QY 61 LDQOSIVHIVQRPKRGEMNATGGDDPRNAGGGEREPQSLTRVDLSSSVLPBGDSVGLA 120
DB 61 LDQOSIVHIVQRPKRGEMNATGGDDPRNAGGGEREPQSLTRVDLSSSVLPBGDSVGLA 120
QY 121 VILHTRSRKDSPPAGSPAGRSIYNSFYVCKGPCQVRQPKLRVQCSCTROATLLTQGP 180
DB 121 VILHTRSRKDSPPAGSPAGRSIYNSFYVCKGPCQVRQPKLRVQCSCTROATLLTQGP 180
QY 181 SCMDVLI.PNRMSGSCQSPHCPTSAEFFKCGAHPSTDKETPVALLIATNSRNITCIT 240
DB 181 SCMDVLI.PNRMSGSCQSPHCPTSAEFFKCGAHPSTDKETPVALLIATNSRNITCIT 240
QY 241 CTDVSPVLVFOCNSRHVYCLDCFHLVCTYRLNDRQFVNDPOLGYSLPVACCPNSLIE 300
DB 241 CTDVSPVLVFOCNSRHVYCLDCFHLVCTYRLNDRQFVNDPOLGYSLPVACCPNSLIE 300
QY 213 CTDVSPVLVFOCNSRHVYCLDCFHLVCTYRLNDRQFVNDPOLGYSLPVACCPNSLIE 272
DB 213 CTDVSPVLVFOCNSRHVYCLDCFHLVCTYRLNDRQFVNDPOLGYSLPVACCPNSLIE 272
QY 301 LHHFRILGEOYNRVQOYGAEBCVLOMGVLCPRPGCAGLLPEPPORVYTCGEGNGLGC 360
DB 273 LHHFRILGEOYNRVQOYGAEBCVLOMGVLCPRPGCAGLLPEPPORVYTCGEGNGLGC 332
QY 361 GFAPCRCKEAYHBCSAVFASGTTTQAYRVDERAABQARWEAASKETIKKTKPCPR 420
DB 333 GFAPCRCKEAYHBCSAVFASGTTTQAYRVDERAABQARWEAASKETIKKTKPCPR 392
QY 421 CHVPEKNGGCMHMKCPQPCRLWCWNGCCEMNRVCMGDHMFV 465
DB 393 CHVPEKNGGCMHMKCPQPCRLWCWNGCCEMNRVCMGDHMFV 437
```

RESULT 5
US-09-949-016-6854
; Sequence 6854, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6854
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6854

Query Match 65.4%; Score 1696.5; DB 4; Length 316;
Best Local Similarity 68.0%; Pred. No. 4, 2e-161;
Matches 316; Conservative 0; Mismatches 0; Indels 149; Gaps 1;

```
QY 1 MIVFVRPNSSHGFPVEVDSITSIFOLKEVYAKRGVPADQLAVIFAGKELRNDWTQVQND 60
DB 1 MIVFVRPNSSHGFPVEVDSITSIFOLKEVYAKRGVPADQLAVIFAGKELRNDWTQVQND 57
QY 61 LDQOSIVHIVQRPKRGEMNATGGDDPRNAGGGEREPQSLTRVDLSSSVLPBGDSVGLA 120
DB 58 ----- 57
QY 121 VILHTRSRKDSPPAGSPAGRSIYNSFYVCKGPCQVRQPKLRVQCSCTROATLLTQGP 180
DB 58 ----- 57
QY 181 SCMDVLI.PNRMSGSCQSPHCPTSAEFFKCGAHPSTDKETPVALLIATNSRNITCIT 240
DB 58 ----- 57
QY 241 CTDVSPVLVFOCNSRHVYCLDCFHLVCTYRLNDRQFVNDPOLGYSLPVACCPNSLIE 300
DB 92 CTDVSPVLVFOCNSRHVYCLDCFHLVCTYRLNDRQFVNDPOLGYSLPVACCPNSLIE 151
QY 301 LHHFRILGEOYNRVQOYGAEBCVLOMGVLCPRPGCAGLLPEPPORVYTCGEGNGLGC 360
DB 152 LHHFRILGEOYNRVQOYGAEBCVLOMGVLCPRPGCAGLLPEPPORVYTCGEGNGLGC 211
QY 361 GFAPCRCKEAYHBCSAVFASGTTTQAYRVDERAABQARWEAASKETIKKTKPCPR 420
DB 212 GFAPCRCKEAYHBCSAVFASGTTTQAYRVDERAABQARWEAASKETIKKTKPCPR 271
QY 421 CHVPEKNGGCMHMKCPQPCRLWCWNGCCEMNRVCMGDHMFV 465
DB 272 CHVPEKNGGCMHMKCPQPCRLWCWNGCCEMNRVCMGDHMFV 316
```

RESULT 6
US-09-270-767-32183
; Sequence 32183, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32183
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-32183

Query Match 14.3%; Score 370.5; DB 4; Length 117;
Best Local Similarity 53.3%; Pred. No. 3, 8e-29;
Matches 64; Conservative 11; Mismatches 40; Indels 5; Gaps 2;

```
QY 345 PDQRVTCGCGNGLGGFAFCRCKEAYHBCG-SAVFASGTTTQAYRVDERAABQARW 403
DB 1 PDQRVTCCKN---GGGVFCRNCLOGYHIGCLBGTGASATNSCEYVDPRAAEARN 56
QY 404 EAASKETIKKTKPCPRCHVPEKNGGCMHMKCPQPCRLWCWNGCCEMNRVCMGDHMF 463
DB 57 DEASNTVITKTPCKKCTPTERDGGCMHMTCTAAGCGFECWVCQTETMTDCKGAMWF 116
```

RESULT 7
US-09-914-259-13
; Sequence 13, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES

FILE REFERENCE: 8471-010-999
 CURRENT APPLICATION NUMBER: US/09/914,259
 CURRENT FILING DATE: 2000-11-21
 NUMBER OF SEQ ID NOS: 180
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO: 13
 LENGTH: 503
 TYPE: PRT
 ORGANISM: Drosophila melanogaster
 US-09-914-259-13

Query Match 9.3%; Score 240.5; DB 4; Length 503;
 Best Local Similarity 27.8%; Pred. No. 3e-15;
 Matches 72; Conservative 42; Mismatches 111; Indels 37; Gaps 13;

QY 209 FPKGAHPTSDKEPVALHLIATNSRITITCTDVRSVPLVQCNRSRVTCLDCEHLYC 268
 DB 105 FFKC-AHVPNPAVTEIKQKTSRSCCEICFSQLPDSNAGLECGHRCFPCMHLYL 163
 QY 269 VTRLANDRQFVHDQLGSLFCAV-GCPNSLIKELHFRILGEOYN-RYQYGADECYLO 326
 DB 164 STK-----IVAGLQOTISCAHGC-DILVDVTVANLVTDARVAVYQQLITNSFV-E 215
 QY 327 MGVLT-CRRPGCG-AGLPEPDQKVTGCGNGLGCGAFRCCKEAYHEG-ECSAVFE 382
 DB 216 CNGLRKCPSTVCTYAVKPYAPRRVCK-----CGHVECFACGENMHPVKCRWL-- 267
 QY 383 ASGTTQAVRVDBRAEQARWEASKEITIKTTKPCPRCHVPEKNGGCMHKCPQPCR 442
 DB 268 -----KKWIKKCDSDSETSNMTAAN-----TKCPRCSVTIRKDGCMHVKCNONCK 315
 QY 443 LEMCNCGCENNRVCMDHMPD 464
 DB 316 NEFCWVCIGSWEP-HGSSWYN 335

RESULT 8
 US-09-354-221-2
 Sequence 2, Application US/09354221
 Patent No. 6699714
 GENERAL INFORMATION:
 APPLICANT: Chang, Chawmhang
 TITLE OF INVENTION: Androgen Receptor Coactivators
 FILE REFERENCE: 920920.90011
 CURRENT APPLICATION NUMBER: US/09/354,221
 CURRENT FILING DATE: 1999-07-15
 EARLIER APPLICATION NUMBER: 60/100,243
 EARLIER FILING DATE: 1998-09-14
 NUMBER OF SEQ ID NOS: 12
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 2
 LENGTH: 474
 TYPE: PRT
 ORGANISM: Homo sapien
 US-09-354-221-2

Query Match 8.6%; Score 222; DB 4; Length 474;
 Best Local Similarity 27.8%; Pred. No. 2e-13;
 Matches 71; Conservative 27; Mismatches 97; Indels 60; Gaps 13;

QY 232 NSRNTC--ITCTDVSPVLV-F-QCNSRHVCLDCEHLYCYTRLANDRQFVHDQLGSLP 288
 DB 214 NSKLFCSICFCEKLSCECMYFLEC--RHVYCKAKLQDFEIQIRDO-----VQ 261
 QY 289 CVAACP-----SLIKELHFRILGEOYNRYQYGADECYLOMGVLT-CRRPGC 338
 DB 262 CL-NCEPKCPVATPGQVKEL-----VEALFARYDRLLQSSLDLMDADVVCPRPCQ 315
 QY 339 AGLLPEPDQKVTGCGNGLGCGAFRCCKEAYH-BEGCSAVFEASGTTQAY----- 391
 DB 316 LPVMEPG-----CTMGICSSCNFAFCTLCRLTYHGVSFCKVTAEKMLDRLNEYLQADEA 370
 QY 392 -----RVDERAAEQARWEASKEITIKTTKPCPRCHVPEKNGGCMHKCPQPCRLE 444

DB 371 NKRLDQRYGKRVYQKALEMESKEWLEKNSKSCPCGTPLEKLDGCMKWTG--TGCWY 428
 QY 445 WCMNCGCENNRVCWG 459
 DB 429 FCW-----ICWG 435

RESULT 9
 US-09-949-016-10111

Sequence 10111, Application US/09949016
 Patent No. 6812339
 GENERAL INFORMATION:
 APPLICANT: VENTER, J. Craig et al.
 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 FILE REFERENCE: CL001307
 CURRENT APPLICATION NUMBER: US/09/949,016
 CURRENT FILING DATE: 2000-04-14
 PRIOR APPLICATION NUMBER: 60/241,755
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/237,768
 PRIOR FILING DATE: 2000-10-03
 PRIOR APPLICATION NUMBER: 60/231,498
 PRIOR FILING DATE: 2000-09-08
 NUMBER OF SEQ ID NOS: 207012
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO: 10111
 LENGTH: 487
 TYPE: PRT
 ORGANISM: Human
 US-09-949-016-10111

Query Match 8.6%; Score 222; DB 4; Length 487;
 Best Local Similarity 27.8%; Pred. No. 2e-13;
 Matches 71; Conservative 27; Mismatches 97; Indels 60; Gaps 13;

QY 232 NSRNTC--ITCTDVSPVLV-F-QCNSRHVCLDCEHLYCYTRLANDRQFVHDQLGSLP 288
 DB 227 NSKLFCSICFCEKLSCECMYFLEC--RHVYCKAKLQDFEIQIRDO-----VQ 274
 QY 289 CVAACP-----SLIKELHFRILGEOYNRYQYGADECYLOMGVLT-CRRPGC 338
 DB 275 CL-NCEPKCPVATPGQVKEL-----VEALFARYDRLLQSSLDLMDADVVCPRPCQ 328
 QY 339 AGLLPEPDQKVTGCGNGLGCGAFRCCKEAYH-BEGCSAVFEASGTTQAY----- 391
 DB 329 LPVMEPG-----CTMGICSSCNFAFCTLCRLTYHGVSFCKVTAEKMLDRLNEYLQADEA 383
 QY 392 -----RVDERAAEQARWEASKEITIKTTKPCPRCHVPEKNGGCMHKCPQPCRLE 444
 DB 384 NKRLDQRYGKRVYQKALEMESKEWLEKNSKSCPCGTPLEKLDGCMKWTG--TGCWY 441
 QY 445 WCMNCGCENNRVCWG 459
 DB 442 FCW-----ICWG 448

RESULT 10

US-09-248-796A-15410
 Sequence 15410, Application US/09248796A
 Patent No. 6747137
 GENERAL INFORMATION:
 APPLICANT: Kelch Weinstein et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
 FILE REFERENCE: 107196.132
 CURRENT APPLICATION NUMBER: US/09/248,796A
 CURRENT FILING DATE: 1999-02-12
 PRIOR APPLICATION NUMBER: US 60/074,725
 PRIOR FILING DATE: 1998-02-13
 PRIOR APPLICATION NUMBER: US 60/096,409
 PRIOR FILING DATE: 1998-08-13

NUMBER OF SEQ ID NOS: 28208
 SEQ ID NO 15410
 LENGTH: 506
 TYPE: prt
 ORGANISM: Candida albicans
 US-09-248-796A-15410

Query Match 7.4%; Score 191; DB 4; Length 506;
 Best Local Similarity 23.6%; Pred. No. 2,7e-10;
 Matches 79; Conservative 48; Mismatches 118; Indels 90; Gaps 17;

QY 166 CSTCQATITLTGSGCWDVILPNMSEGCSPHCPTSAEFPKCGAHPISDKETPVA 225
 DB 207 CQIC-QKTL---KGLHC-----TKPDCGHVFNCNCLAEYFESG----- 241
 QY 226 LHLITNSNITCTITDVRSPVLVFNQNSRVHVICDCEPHLYCVTRLNROFVHPDQLGY 285
 DB 242 -----IESGDIDKVNCPD-----FECTKKYVDKKKKF-MELETVVHNDKKVKDMLNQL 288
 QY 286 SLPCVAGCNSLIKELHFRILGSEBOYNYO--QYGAECVLQMGVLPFRPGCAGLLP 343
 DB 289 LIPST---PLAMLTKLKSPSLVDRYNLPFKSQYEWIGNLPLNRLVKCPRGCDDEVIR 345
 QY 344 EP-DQRKVTCEGNGLGCGPAFCRECKEAYHE-----GECSAVFEASGTTQAY--- 391
 DB 346 EDIDELVYCP-----KCKYAFNCNCRKSYHARFKLCLKINENDKYLGIPIEDLNAVPL 400
 QY 392 -----RVDERAEOARWEAASKETI--KTTKPCPRCHVPEKNGCGM 432
 DB 401 PPDSDYDKKIMNAKYGRKRID-PAIEEYQMDOLFQMMRERKTVVQCPGCSVATEKSEGCN 459
 QY 433 HMKCPQPCRLMCMNCGCEMNRVCMG---DHWFD 464
 DB 460 KMKC-SLCKTDFCFNGSK-----IGNNHDFVD 487

RESULT 11
 US-08-398-008A-2
 Sequence 2, Application US/08398008A
 Patent No. 5665388
 GENERAL INFORMATION:
 APPLICANT: Kornbluth, Jacki
 TITLE OF INVENTION: DNA Encoding Natural Killer Lytic Associated
 TITLE OF INVENTION: Protein
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Glibbrech & Adler, P.C.
 STREET: 8011 Candle Lane
 CITY: Houston
 STATE: Texas
 COUNTRY: USA
 ZIP: 77071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
 COMPUTER: MACINTOSH IICI
 OPERATING SYSTEM: Macintosh
 SOFTWARE: Microsoft Word 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/398,008A
 FILING DATE: March 2, 1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/126,501
 FILING DATE: 24-SEP-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Adler, Dr. Benjamin Aaron
 REGISTRATION NUMBER: 35,423
 REFERENCE/DOCKET NUMBER: D5705CIP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (713) 777-2321
 TELEFAX: (713) 777-6908
 TELEX:
 INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
 LENGTH: 587
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: Protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-398-008A-2

Query Match 6.4%; Score 167; DB 1; Length 587;
 Best Local Similarity 24.2%; Pred. No. 8.4e-08;
 Matches 80; Conservative 24; Mismatches 114; Indels 112; Gaps 20;

QY 178 QGPGCWDVILIPNRMSGECSPHCPTSAEF-FPKCGAHPISDKETPVALLIATNSRNI 236
 DB 73 QGPP-----PEALPAE-----PAAEAEAAAAAGAPGDDDE--AAEGGPGGEV 117
 QY 237 TCITC---TDVRSPLVFNQNSRVHVICDCEPHLY-----CVTRLND-- 274
 DB 118 ECPCLCVRLPERRAPRL---SCPHRSRCDLRHLYRLISRSRVISCPGCSERLNPMD 174
 QY 275 -RQFVHPDQGYSLPCVAGCNSLIKELHFRILGSEBOYNYOQYGAECVLQMGVLCF 333
 DB 175 IRLILADPPL-----MHKY-----EEFVRRYLASDPDC-----RWCP 207
 QY 334 RPGCGAGLL-----PEPDQRKVTCEGNGLGCGPAFCRECKEAYHEGECSAVFEASGT 386
 DB 208 APDCGAVLAAYCASC-----KLTCERE--GCOTEFYHKQIWHNPQTCM--ARQO 257
 QY 387 TQAVRVDERAEOARWEAASKETIKTTKPCPRCHVPEK--NGSCMAMKCPQPCRL 444
 DB 258 RAQTLARVTKHTSGLSY--GQESGPDILKPCPRCSAYIIRKNDSCNMTC--AVCGGE 312
 QY 445 WCMNC-----GC-----ENMR 455
 DB 313 FCLMCKEISDLHYLSPSGCTFWGKKPWR 342

RESULT 12
 US-08-893-333-2
 Sequence 2, Application US/08893333A
 Patent No. 5981705
 GENERAL INFORMATION:
 APPLICANT: Kornbluth, Jacki
 TITLE OF INVENTION: DNA Encoding Natural Killer Lytic Associated Protein
 FILE REFERENCE: D5705CIP/D
 CURRENT APPLICATION NUMBER: US/08/893,333A
 CURRENT FILING DATE: 1997-07-16
 NUMBER OF SEQ ID NOS: 17
 SEQ ID NO 2
 LENGTH: 587
 TYPE: prt
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: Amino acid sequence of Natural Killer Lytic
 OTHER INFORMATION: Associated Protein encoded by nucleotides
 OTHER INFORMATION: 190 to 1953 of Sequence ID. No. 5981705
 Patent No. 5981705
 US-08-893-333-2

Query Match 6.4%; Score 167; DB 2; Length 587;
 Best Local Similarity 24.2%; Pred. No. 8.4e-08;
 Matches 80; Conservative 24; Mismatches 114; Indels 112; Gaps 20;

QY 178 QGPGCWDVILIPNRMSGECSPHCPTSAEF-FPKCGAHPISDKETPVALLIATNSRNI 236
 DB 73 QGPP-----PEALPAE-----PAAEAEAAAAAGAPGDDDE--AAEGGPGGEV 117
 QY 237 TCITC---TDVRSPLVFNQNSRVHVICDCEPHLY-----CVTRLND-- 274
 DB 118 ECPCLCVRLPERRAPRL---SCPHRSRCDLRHLYRLISRSRVISCPGCSERLNPMD 174

QY 275 -RQFVHDPLQVLSLFCVAGCPSNLIKELHFRILIEEOYNYRQOYGABCVLQMGVLCF 333
DB 175 IRLIADPL-----MHXY-----EEFIRRIYASPPD-----RWCP 207
QY 334 RFGCAGL-----PEPDARKYTCGAGLGGCFPFCECKEAYHEGCSAVFEASGT 386
DB 208 APDGGYAVIANGCASC-----KLTCEKE-----GQTEFCHCQIWHNPQTCM--ARQ 257
QY 387 TTAQVYDERAAEQARWEASKETIKTKTPRCCHVPEK--NGGCMKCPQPCRL 444
DB 258 RAQTLARVTKHTSGLSY---GQSGPDDIKPCRCGAYIIKMDGSCNMHTC--AVCGCE 312
QY 445 WCNAC-----GC-----EMNR 455
DB 313 FCWLCKEISDILHLSFGCTFWGKKRWSR 342

RESULT 13
US-09-538-092-598
; Sequence 598, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Gluc, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; NUMBER OF SEQ ID NOS: 2000-02-01
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 598
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YNL068W
US-09-538-092-598

Query Match
Best Local Similarity 5.7%; Score 149; DB 4; Length 464;
Matches 59; Conservative 33; Mismatches 105; Indels 78; Gaps 13;
QY 235 NITCITCTDVRSPLYVFO-----CNSRHVCLDPCFHLVCVTRINDROF--VHDPLQVLS- 286
DB 177 NYHCICMEHEKGYRMTKLPCEANAVEHYLCRCGCAKSYFTAMIQENRISVACPCCEYKE 236
QY 287 -----LPCVACGPNLSIKELHFRILIEEOYNYRQOYGABCVLQMGVLCF 329
DB 237 LKLEDFKSYKMKLKFPLI---PVSPLKE-----VIDTELCEYKEKFFVQAATRIISK 288
QY 330 -----VLCPRGCGAGLPEP-DQKRYTCGAGLGGCFPFCECKEAYH--BEGCSAV 380
DB 289 YCPYACVTCRR--CDSWCTKEDLDAMIQO-----KGFVFCGFCDLHAMGYNNKCGKK 341
QY 381 FEASGTTTQAY-----RYDERAAEQARWEASKETIKTKT 415
DB 342 VSLSTDIIEYLDVTVTSYERKRLKAKYGRVLEBYNDVIAEKMDLAIKE--GSNL 399
QY 416 KCPRCCHVPEKNGGCMKCPQPCRLIEWCMNGCEMNRVCMGDH 461
DB 400 QRCPCYKVERSEGCNKKRC--EVCGLTFCFTICG 432

RESULT 14
US-09-248-796A-16471
; Sequence 16471, Application US/09248796A
; Patent No. 6747137

; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; NUMBER OF SEQ ID NOS: 1998-08-13
; SEQ ID NO 16471
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16471

Query Match
Best Local Similarity 5.4%; Score 141; DB 4; Length 328;
Matches 32; Conservative 15; Mismatches 37; Indels 22; Gaps 5;
QY 358 LCGG--FAFCRECKEAYHEGCSAVFEASGTTQAYVDERAAEQARWEASKETIKTKT 415
DB 58 ICGSHHEFCNCKENTL-PCPC-----WVSRRWIKKCNDSDETAHWIDANT 104
QY 416 KCPRCCHVPEKNGGCMKCPQPCRLIEWCMNGCEMNRVCMGDH 461
DB 105 HSCPCCYSSIEKNGCNHMT--QKCYEFCWCWCLKDWS-----DH 143

RESULT 15
US-08-854-764-2
; Sequence 2, Application US/08854764
; Patent No. 6103500
; GENERAL INFORMATION:
; APPLICANT: Innes, Michael
; APPLICANT: Creasey, Abia
; TITLE OF INVENTION: Production of Tissue Factor Pathway
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
ADDRESS: Chilton Corporation
STREET: 4560 Horton St.
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
; COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.308
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,764
FILING DATE: 12-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/286,530
FILING DATE: 05-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Savereide, Paul B.
REGISTRATION NUMBER: 36,914
REFERENCE/DOCKET NUMBER: 0991.001
TELEPHONE: 510-655-3542
TELEPHONE: 510-601-2585
; INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPER: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-854-764-2

Query Match 5.1%; Score 133; DB 3; Length 352;
Best Local Similarity 18.3%; Pred. No. 0.0001;
Matches 70; Conservative 56; Mismatches 116; Indels 140; Gaps 16;

```
QY      1 MIVFVFNSSHGFPVEVDSDTSI FOLKEVAKQGVADQLRVIFAGKELRNDWTVQNC 60
      1 MGI FVKTLTGKITLEVESDITDNVSKIQDEKGI PPQQRILIFAGKCLEDRITLSDYN 60
DB
QY      61 LDQGSIVHIVQRPWKKGEMNATGGDDPKNAAAGCEREPQSILTRVDLS SVPGDSVGLA 120
      61 IQKESTLHLVLR-----LRGGDSEE-----DEHTIITDTLPLPLK----- 97
DB
QY      121 VILHT--DSRKDSPAGSPAGRSIYNSF-----VYV-----CKGPCORV 157
      98 --MHSFCAFKADDPCKAIMKRRFFNI FTROCEBFYIGCEGNQNRFEHLECKKKCTRD 155
DB
QY      158 QPGKLRVOCSTCRQATLTLTGSPSCWDVLI PNRMSEGCSPH-CPGTGAEEFFK----- 211
      156 NANRI-----IKTTLQOEKDPFCLE-----EDPGICRGYITRYFYNNQTKQ 197
DB
QY      212 -----CGAHPISDKETPVALLHLIATNSRNITCITCTDVS 246
      198 CERFHYGGCLGNMNNPETLECKNICEDGPNGPOVDNYGTOLNAVNN---SLTPQSTRV 253
DB
QY      247 PVLVPQCNSRHVICIDCFH--LYCVT-----RLNDRQFVHDPQLGYSLP----- 288
      254 PSL-PE-----FHGPSMCLTPADRGLCRANENRFYNSVIGKCRPFKYSGCCGN 301
DB
QY      289 -----CVAGCPNSLIKEL 301
      302 ENNFTSKOECLRAKCKKGFIORI 323
```

Search completed: February 23, 2005, 14:01:03
Job time : 27.8801 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM proteain - proteain search, using sw model

Run on: February 23, 2005, 13:57:26 ; Search time 73.0:13 Seconds
(without alignments)
2084.158 Million cell updates/sec

Title: US-10-622-817-8
Perfect score: 2596
Sequence: 1 MIVFVRFNSSHGPFVEVSDS.....CINCCGCEMNRVCMDGDFV 465

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|--------------------|
| 1 | 2596 | 100.0 | 465 | US-10-473-226-2 | Sequence 2, Appl1 |
| 2 | 2596 | 100.0 | 465 | US-10-776-604-2 | Sequence 2, Appl1 |
| 3 | 2586 | 100.0 | 465 | US-10-839-688-9 | Sequence 9, Appl1 |
| 4 | 2584 | 99.5 | 465 | US-10-473-226-4 | Sequence 4, Appl1 |
| 5 | 2405 | 93.6 | 437 | US-10-776-604-4 | Sequence 4, Appl1 |
| 6 | 870 | 33.5 | 156 | US-09-785-548-4 | Sequence 4, Appl1 |
| 7 | 544.5 | 21.0 | 386 | US-10-239-249-2 | Sequence 10, Appl1 |
| 8 | 373 | 14.4 | 64 | US-10-313-203-10 | Sequence 17, Appl1 |
| 9 | 323 | 12.4 | 56 | US-10-313-203-17 | Sequence 17, Appl1 |
| 10 | 263 | 10.1 | 46 | US-09-864-761-36750 | Sequence 36750, A |
| 11 | 246 | 9.5 | 525 | US-10-437-963-199787 | Sequence 55407, A |
| 12 | 242.5 | 9.3 | 604 | US-10-425-114-55407 | Sequence 13, Appl1 |
| 13 | 240.5 | 9.3 | 503 | US-10-080-608A-13 | Sequence 13, Appl1 |

| | | | | | |
|----|-------|-----|------|----------------------|--------------------|
| 14 | 240.5 | 9.3 | 503 | US-10-370-685-102 | Sequence 102, App |
| 15 | 233.5 | 9.0 | 585 | US-10-424-599-147673 | Sequence 147673, A |
| 16 | 224.5 | 8.6 | 648 | US-10-425-114-63379 | Sequence 63379, A |
| 17 | 222 | 8.6 | 474 | US-09-808-387-42 | Sequence 42, Appl1 |
| 18 | 222 | 8.6 | 474 | US-10-442-754-2 | Sequence 2, Appl1 |
| 19 | 222 | 8.6 | 474 | US-10-408-765A-572 | Sequence 572, App |
| 20 | 222 | 8.6 | 477 | US-10-264-049-3019 | Sequence 3019, Ap |
| 21 | 219.5 | 8.5 | 1753 | US-10-146-473-44 | Sequence 44, Appl1 |
| 22 | 219.5 | 8.5 | 1753 | US-10-276-774-1824 | Sequence 1824, Ap |
| 23 | 219.5 | 8.5 | 2517 | US-10-313-203-2 | Sequence 2, Appl1 |
| 24 | 215 | 8.3 | 32 | US-10-313-203-18 | Sequence 18, Appl1 |
| 25 | 207.5 | 8.0 | 623 | US-10-437-963-148905 | Sequence 148905, A |
| 26 | 206.5 | 8.0 | 452 | US-10-425-114-38443 | Sequence 38443, A |
| 27 | 205.5 | 7.9 | 449 | US-10-425-114-39657 | Sequence 39657, A |
| 28 | 204 | 7.9 | 624 | US-10-374-780A-1727 | Sequence 1727, Ap |
| 29 | 204 | 7.9 | 624 | US-10-437-963-140793 | Sequence 140793, A |
| 30 | 202.5 | 7.8 | 289 | US-10-437-963-133385 | Sequence 133385, A |
| 31 | 201 | 7.7 | 333 | US-10-424-599-217796 | Sequence 217796, A |
| 32 | 200 | 7.7 | 593 | US-10-225-066A-854 | Sequence 854, App |
| 33 | 200 | 7.7 | 593 | US-10-225-066A-854 | Sequence 56, Appl1 |
| 34 | 200 | 7.7 | 593 | US-10-225-066A-854 | Sequence 330, App |
| 35 | 195.5 | 7.5 | 541 | US-10-424-599-265215 | Sequence 265215, A |
| 36 | 193 | 7.4 | 606 | US-10-437-963-111378 | Sequence 111378, A |
| 37 | 192 | 7.4 | 354 | US-10-425-114-56248 | Sequence 56248, A |
| 38 | 191.5 | 7.4 | 540 | US-10-437-963-168788 | Sequence 168788, A |
| 39 | 191 | 7.4 | 338 | US-10-425-114-56467 | Sequence 56467, A |
| 40 | 184 | 7.1 | 583 | US-10-437-963-124556 | Sequence 124556, A |
| 41 | 182.5 | 7.0 | 522 | US-10-437-963-184651 | Sequence 184651, A |
| 42 | 180 | 6.9 | 250 | US-10-437-963-152161 | Sequence 152161, A |
| 43 | 180 | 6.9 | 394 | US-10-425-114-54725 | Sequence 54725, A |
| 44 | 180 | 6.9 | 557 | US-10-424-599-212592 | Sequence 212592, A |
| 45 | 177.5 | 6.8 | 396 | US-10-425-114-40231 | Sequence 40231, A |

ALIGNMENTS

RESULT 1
US-10-473-226-2
Sequence 2, Application US/10473226
GENERAL INFORMATION:
APPLICANT: NGENE A/S
TITLE OF INVENTION: Means for inhibiting proteolytical processing of Parkin
FILE REFERENCE: 506-204-WO
CURRENT FILING DATE: 2003-09-29
PRIOR APPLICATION NUMBER: US/10/473,226
PRIOR FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: DK PA 2001 00525
PRIOR FILING DATE: 2001-04-03
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 465
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: mat_peptide
LOCATION: (1)..(465)
OTHER INFORMATION: Native Parkin
US-10-473-226-2

Query Match 100.0% Score 2596; DB 16; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.2e+220;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MIVFVRFNSSHGPFVEVSDSIFOLKEVYVARGQVPAQDLVIFAGKELRMDWYQNC 60
DB 1 MIVFVRFNSSHGPFVEVSDSIFOLKEVYVARGQVPAQDLVIFAGKELRMDWYQNC 60
QY 61 LDQOSIVHVRPMKRGEMNATGDDPRNNAAGCEREPQSLTRVDLSSSVLPGDSVGLA 120

Db 61 LDQOSIVHIYQRPWRKQGEENATGDDPRNAAGGCEREPOSILTRVDSLSSVLPAGDSVGLA 120
QY 121 VILHTDSRKDSPAGSPAGRSIYNSFYVYCKGPCQVOPGKLRVOCSTCRQATLTLTQGP 180
QY 121 VILHTDSRKDSPAGSPAGRSIYNSFYVYCKGPCQVOPGKLRVOCSTCRQATLTLTQGP 180
QY 181 SCMDVLLIPNRMSGECQSPHCPGTSAPFPKCGAHTSDKETPVALHLIATNSRNTICIT 240
Db 181 SCMDVLLIPNRMSGECQSPHCPGTSAPFPKCGAHTSDKETPVALHLIATNSRNTICIT 240
QY 241 CTDVASPVLVFOQNSRHVILCLDFHLYCVTRLRNDQFVHDPOLGYSLPCVACGPNLSIKE 300
Db 241 CTDVASPVLVFOQNSRHVILCLDFHLYCVTRLRNDQFVHDPOLGYSLPCVACGPNLSIKE 300
QY 301 LHHFRILGEEQYNRYYQYGAEBECVLQMGVLCPRPGCAGLLPEPDQRKVTCEGNGIGC 360
Db 301 LHHFRILGEEQYNRYYQYGAEBECVLQMGVLCPRPGCAGLLPEPDQRKVTCEGNGIGC 360
QY 361 GFAPFCECKEAYHEGECSAVFPAASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
Db 361 GFAPFCECKEAYHEGECSAVFPAASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
QY 421 CHVPEKNGGCMHMKCPQPCRLIEWCMNCGEENRVCMGDHMFV 465
Db 421 CHVPEKNGGCMHMKCPQPCRLIEWCMNCGEENRVCMGDHMFV 465

RESULT 2

US-10-776-604-2
; Sequence 2, Application US/10776604
; Publication No. US20050003385A1
; GENERAL INFORMATION:
; APPLICANT: Shimizu, Nobuyoshi
; APPLICANT: Mizuno, Yoshikuni
; TITLE OF INVENTION: DAs or Genes Participating in Parkinson's Disease
; FILE REFERENCE: 0652.2110001
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 09/601,844
; PRIOR FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: PCT/JP99/00545
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: JP 10/27531
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-776-604-2

Query Match 100.0%; Score 2596; DB 16; Length 465;
Best Local Similarity 100.0%; Pred. No. 2,2e-220;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIVFVRNNSHGPFVEVSDTSTIFOLKEVYAKQGVADQLRVIFAGKELRNDWTQVONCD 60
Db 1 MIVFVRNNSHGPFVEVSDTSTIFOLKEVYAKQGVADQLRVIFAGKELRNDWTQVONCD 60
QY 61 LDQOSIVHIYQRPWRKQGEENATGDDPRNAAGGCEREPOSILTRVDSLSSVLPAGDSVGLA 120
Db 61 LDQOSIVHIYQRPWRKQGEENATGDDPRNAAGGCEREPOSILTRVDSLSSVLPAGDSVGLA 120
QY 121 VILHTDSRKDSPAGSPAGRSIYNSFYVYCKGPCQVOPGKLRVOCSTCRQATLTLTQGP 180
Db 121 VILHTDSRKDSPAGSPAGRSIYNSFYVYCKGPCQVOPGKLRVOCSTCRQATLTLTQGP 180
QY 181 SCMDVLLIPNRMSGECQSPHCPGTSAPFPKCGAHTSDKETPVALHLIATNSRNTICIT 240
Db 181 SCMDVLLIPNRMSGECQSPHCPGTSAPFPKCGAHTSDKETPVALHLIATNSRNTICIT 240
QY 241 CTDVASPVLVFOQNSRHVILCLDFHLYCVTRLRNDQFVHDPOLGYSLPCVACGPNLSIKE 300
Db 241 CTDVASPVLVFOQNSRHVILCLDFHLYCVTRLRNDQFVHDPOLGYSLPCVACGPNLSIKE 300

Db 241 CTDVASPVLVFOQNSRHVILCLDFHLYCVTRLRNDQFVHDPOLGYSLPCVACGPNLSIKE 300
QY 301 LHHFRILGEEQYNRYYQYGAEBECVLQMGVLCPRPGCAGLLPEPDQRKVTCEGNGIGC 360
Db 301 LHHFRILGEEQYNRYYQYGAEBECVLQMGVLCPRPGCAGLLPEPDQRKVTCEGNGIGC 360
QY 361 GFAPFCECKEAYHEGECSAVFPAASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
Db 361 GFAPFCECKEAYHEGECSAVFPAASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
QY 421 CHVPEKNGGCMHMKCPQPCRLIEWCMNCGEENRVCMGDHMFV 465
Db 421 CHVPEKNGGCMHMKCPQPCRLIEWCMNCGEENRVCMGDHMFV 465

RESULT 3

US-10-839-688-9
; Sequence 9, Application US/10839688
; Publication No. US20050014173A1
; GENERAL INFORMATION:
; APPLICANT: Farrer, Matthew J.
; TITLE OF INVENTION: PARKINSON'S DISEASE MARKERS
; FILE REFERENCE: 07039-448001
; CURRENT FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: US 60/468,832
; PRIOR FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-839-688-9

Query Match 100.0%; Score 2596; DB 17; Length 465;
Best Local Similarity 100.0%; Pred. No. 2,2e-220;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIVFVRNNSHGPFVEVSDTSTIFOLKEVYAKQGVADQLRVIFAGKELRNDWTQVONCD 60
Db 1 MIVFVRNNSHGPFVEVSDTSTIFOLKEVYAKQGVADQLRVIFAGKELRNDWTQVONCD 60
QY 61 LDQOSIVHIYQRPWRKQGEENATGDDPRNAAGGCEREPOSILTRVDSLSSVLPAGDSVGLA 120
Db 61 LDQOSIVHIYQRPWRKQGEENATGDDPRNAAGGCEREPOSILTRVDSLSSVLPAGDSVGLA 120
QY 121 VILHTDSRKDSPAGSPAGRSIYNSFYVYCKGPCQVOPGKLRVOCSTCRQATLTLTQGP 180
Db 121 VILHTDSRKDSPAGSPAGRSIYNSFYVYCKGPCQVOPGKLRVOCSTCRQATLTLTQGP 180
QY 181 SCMDVLLIPNRMSGECQSPHCPGTSAPFPKCGAHTSDKETPVALHLIATNSRNTICIT 240
Db 181 SCMDVLLIPNRMSGECQSPHCPGTSAPFPKCGAHTSDKETPVALHLIATNSRNTICIT 240
QY 241 CTDVASPVLVFOQNSRHVILCLDFHLYCVTRLRNDQFVHDPOLGYSLPCVACGPNLSIKE 300
Db 241 CTDVASPVLVFOQNSRHVILCLDFHLYCVTRLRNDQFVHDPOLGYSLPCVACGPNLSIKE 300
QY 301 LHHFRILGEEQYNRYYQYGAEBECVLQMGVLCPRPGCAGLLPEPDQRKVTCEGNGIGC 360
Db 301 LHHFRILGEEQYNRYYQYGAEBECVLQMGVLCPRPGCAGLLPEPDQRKVTCEGNGIGC 360
QY 361 GFAPFCECKEAYHEGECSAVFPAASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
Db 361 GFAPFCECKEAYHEGECSAVFPAASGTTTQAYRVDERAAEQARWEAASKETIKKTTKPCPR 420
QY 421 CHVPEKNGGCMHMKCPQPCRLIEWCMNCGEENRVCMGDHMFV 465
Db 421 CHVPEKNGGCMHMKCPQPCRLIEWCMNCGEENRVCMGDHMFV 465

```
RESULT 4
US-10-473-226-4
; Sequence 4, Application US/10473226
; Publication No. US20040198650A1
; GENERAL INFORMATION:
; APPLICANT: NGENE A/S
; TITLE OF INVENTION: Means for inhibiting proteolytical processing of Parkin
; FILE REFERENCE: 506-204-WO
; CURRENT APPLICATION NUMBER: US/10/473,226
; CURRENT FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: DK PA 2001 00525
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: US 60/281,286
; PRIOR FILING DATE: 2001-04-03
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.1
; SEQ ID NO: 4
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mac_peptide
; LOCATION: (1)..(1)
; OTHER INFORMATION: Parkin with a D126E mutation
US-10-473-226-4

Query Match
Best Local Similarity 99.5%; Score 2584; DB 16; Length 465;
Matches 463; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MIVFRRFNSHGFPEVVDSDTSIFOLKEVVAKRQGVADQLRVIFAGKELRNDWTQVQND 60
DB 1 MIVFRRFNSHGFPEVVDSDTSIFOLKEVVAKRQGVADQLRVIFAGKELRNDWTQVQND 60
QY 61 LDQOSIVHIVQRPWRKQGMNATGDDPRNAAAGCCEREPQSLTRVDLSSSVLPQDSVGLA 120
DB 61 LDQOSIVHIVQRPWRKQGMNATGDDPRNAAAGCCEREPQSLTRVDLSSSVLPQDSVGLA 120
QY 121 VILHTDSRSDSPASSPAGRSIYNSFYVCKPCQRPVQGKLRVOCSTORATLTITQGP 180
DB 121 VILHTDSRSDSPASSPAGRSIYNSFYVCKPCQRPVQGKLRVOCSTORATLTITQGP 180
QY 121 VILHTDSRSDSPASSPAGRSIYNSFYVCKPCQRPVQGKLRVOCSTORATLTITQGP 180
DB 121 VILHTDSRSDSPASSPAGRSIYNSFYVCKPCQRPVQGKLRVOCSTORATLTITQGP 180
QY 181 SCMDVLLIPNRMSGECQSPHCPTSAEFPFKCAHPTSDKETPVALLHATNSRNITCIT 240
DB 181 SCMDVLLIPNRMSGECQSPHCPTSAEFPFKCAHPTSDKETPVALLHATNSRNITCIT 240
QY 241 CTDVASPVLVFOCNSRHVILCLDCFHLVYCTRLNDRQFVHDPLQGYSLPCVAGCPNSLIKE 300
DB 241 CTDVASPVLVFOCNSRHVILCLDCFHLVYCTRLNDRQFVHDPLQGYSLPCVAGCPNSLIKE 300
QY 301 LHHFRILGEBEQNRYQOYGAEECVLQMGVLCPRPCGAGLLPEPQRYTCGEGNGLGC 360
DB 301 LHHFRILGEBEQNRYQOYGAEECVLQMGVLCPRPCGAGLLPEPQRYTCGEGNGLGC 360
QY 361 GFAFRCEKEAYHBECSAVFEASGTTTQAYRVEDRAAQAWEAASKETIKKTPCPR 420
DB 361 GFAFRCEKEAYHBECSAVFEASGTTTQAYRVEDRAAQAWEAASKETIKKTPCPR 420
QY 421 CHVPEKNGGCMHMKCPQPCRLIEWCMNCGCEMNRVCMGDHMFV 465
DB 421 CHVPEKNGGCMHMKCPQPCRLIEWCMNCGCEMNRVCMGDHMFV 465

RESULT 5
US-10-776-604-4
; Sequence 4, Application US/10776604
; Publication No. US2005000385A1
; GENERAL INFORMATION:
; APPLICANT: Shimizu, Nobuyoshi
; APPLICANT: Mizuno, Yoshikuni
; TITLE OF INVENTION: DNAs or Genes Participating in Parkinson's Disease
; FILE REFERENCE: 0652.2110001
; CURRENT APPLICATION NUMBER: US/10/776,604
```

```
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 09/601,844
; PRIOR FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: PCT/JP99/00545
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: JP 10/27531
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patentin version 3.1
; SEQ ID NO: 4
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-776-604-4

Query Match
Best Local Similarity 94.0%; Score 2405; DB 16; Length 437;
Matches 437; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

QY 1 MIVFRRFNSHGFPEVVDSDTSIFOLKEVVAKRQGVADQLRVIFAGKELRNDWTQVQND 60
DB 1 MIVFRRFNSHGFPEVVDSDTSIFOLKEVVAKRQGVADQLRVIFAGKELRNDWTQVQND 60
QY 61 LDQOSIVHIVQRPWRKQGMNATGDDPRNAAAGCCEREPQSLTRVDLSSSVLPQDSVGLA 120
DB 61 LDQOSIVHIVQRPWRKQGMNATGDDPRNAAAGCCEREPQSLTRVDLSSSVLPQDSVGLA 120
QY 121 VILHTDSRSDSPASSPAGRSIYNSFYVCKPCQRPVQGKLRVOCSTORATLTITQGP 180
DB 121 VILHTDSRSDSPASSPAGRSIYNSFYVCKPCQRPVQGKLRVOCSTORATLTITQGP 180
QY 121 VILHTDSRSDSPASSPAGRSIYNSFYVCKPCQRPVQGKLRVOCSTORATLTITQGP 178
DB 121 VILHTDSRSDSPASSPAGRSIYNSFYVCKPCQRPVQGKLRVOCSTORATLTITQGP 178
QY 181 SCMDVLLIPNRMSGECQSPHCPTSAEFPFKCAHPTSDKETPVALLHATNSRNITCIT 240
DB 179 -----EFPFKCAHPTSDKETPVALLHATNSRNITCIT 212
QY 241 CTDVASPVLVFOCNSRHVILCLDCFHLVYCTRLNDRQFVHDPLQGYSLPCVAGCPNSLIKE 300
DB 241 CTDVASPVLVFOCNSRHVILCLDCFHLVYCTRLNDRQFVHDPLQGYSLPCVAGCPNSLIKE 272
QY 301 LHHFRILGEBEQNRYQOYGAEECVLQMGVLCPRPCGAGLLPEPQRYTCGEGNGLGC 360
DB 273 LHHFRILGEBEQNRYQOYGAEECVLQMGVLCPRPCGAGLLPEPQRYTCGEGNGLGC 332
QY 361 GFAFRCEKEAYHBECSAVFEASGTTTQAYRVEDRAAQAWEAASKETIKKTPCPR 420
DB 333 GFAFRCEKEAYHBECSAVFEASGTTTQAYRVEDRAAQAWEAASKETIKKTPCPR 392
QY 421 CHVPEKNGGCMHMKCPQPCRLIEWCMNCGCEMNRVCMGDHMFV 465
DB 393 CHVPEKNGGCMHMKCPQPCRLIEWCMNCGCEMNRVCMGDHMFV 437

RESULT 6
US-09-785-548-4
; Sequence 4, Application US/09785548
; Patent No. US20020155577A1
; GENERAL INFORMATION:
; APPLICANT: AVENTIS PHARMACEUTICALS, INC.
; TITLE OF INVENTION: COMPOSITIONS THAT CAN BE USED FOR REGULATING THE ACTIVITY OF PARK
; FILE REFERENCE: ST00005
; CURRENT APPLICATION NUMBER: US/09/785,548
; CURRENT FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO: 4
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-785-548-4

Query Match
Best Local Similarity 93.5%; Score 870; DB 9; Length 156;
Matches 155; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

QY      135 GSPAGRSIYNFYVYCKGPGCORVOPGKLRVOCSTCRQATLTLTGPGSCMDVLIIPNMSG 194
      |||
Db      1 GSPAGRSIYNFYVYCKGPGCORVOPGKLRVOCSTCRQATLTLTGPGSCMDVLIIPNMSG 60
QY      195 ECOSPHCGTSAEFPFKCGAHPSTDKETPVALLIATNSRNITCITCTVRSPLVAFQCN 254
      |||
Db      61 ECOSPHCGTSAEFPFKCGAHPSTDKETSVALLIATNSRNITCITCTVRSPLVAFQCN 120
QY      255 SRHVICLDCEFLYCVTRLNDROPVHDPOLGYSLEPCV 290
      |||
Db      121 SRHVICLDCEFLYCVTRLNDROPVHDPOLGYSLEPCV 156

RESULT 7
US-10-239-249-2
; Sequence 2, Application US/10239249.
; Publication No. US20030177507A1
; GENERAL INFORMATION:
; APPLICANT: HONER, MARIUS
; APPLICANT: LINK, MOLEFANG
; APPLICANT: BAUMEISTER, RALF
; TITLE OF INVENTION: NEMATODES AS MODEL ORGANISMS FOR INVESTIGATING
; TITLE OF INVENTION: NEURODEGENERATIVE DISEASES AND, IN PARTICULAR, PARKINSON'S
; TITLE OF INVENTION: DISEASE, USES AND METHODS FOR FINDING SUBSTANCES AND
; TITLE OF INVENTION: GENES WHICH CAN BE USED IN TREATING SUCH DISEASES, AND
; FILE REFERENCE: 02481.1804-00000
; CURRENT APPLICATION NUMBER: US/10/239,249
; PRIOR APPLICATION NUMBER: PCT/EP01/03214
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 100 14 109.9
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-239-249-2

Query Match      21.0%; Score 544.5; DB 14; Length 386;
Best Local Similarity 28.9%; Pred. No. 2.1e-39;
Matches 129; Conservative 59; Mismatches 166; Indels 93; Gaps 13;

QY      22 SIFOLKEVAVAKRGVADQLRVIFAGKELRNDVTVOQCDLDOOSIYHIVQRMKGQEWN 81
      |||
Db      29 NIEDLTKDVEKLTETPSDELEVFCCKLSTIMRDLSTRTQIMLRKFNHENG 88
QY      82 ATCGDDPRAAAGCCEREPOSLTRVDLSSVLPDGSVGLAVIILHDSRKDSPAGSPAGRS 141
      |||
Db      89 AT-----TAKITTD-----S 99
QY      142 IYNSFYVYCKGPGCORVOPGKLRVOCSTCRQATLTLTGPGSCMDVLIIPNMSGECOSPHC 201
      |||
Db      100 IIGSFYVWGM-KCDVYKRGKLRVYCKCKSTVLVAKSEQNSDVLKSRRLPAVCEECT 158
QY      202 PGTSAEFPFKCGAHPSTDKETPVALLIATNSRNITCITCTVRSPLVAFQCN SRHVICL 261
      |||
Db      159 PGTSAEFPFKC-----LACNDPAALLTHVGNMOMECCKDCKEAVITDGLGN-HITCQ 212
QY      262 DEFHLYCTRLNDROPVHDPOLGYSLEPC-VAGCPNLSLKEHLHFRILBEOYNRYQYGA 320
      |||
Db      213 FCFRDIYLSQERFVFVQPPHGFIFCYPG-C-NRVQDVHFRIMQTSYSEYQKAT 271
QY      321 BECV-LQMGVYLCPRPCCAGLLPAP--DORKVTEGGNGLGGPAPFRECCEAVHBECC 377
      |||
Db      272 ERLIANDKGVYCPNVSCGSPFMEPYDDGSGCP-----DCEFSFKKC-----PERNC 322
QY      378 SAVFASGTTTOAYVVDRAAEQARWEAASKETIKTKTPCPRCHVPEKNCGCMKPCP 437
      |||
Db      323 VQSHEDLTRLT-----TIDATRRCPCKGVATERNGGCAHHC- 360

```

```

QY      438 QPOCRLEMCNGCCENRNVCMGDHWD 464
      |||
Db      361 -TSCGMWCFCCKTEWKECQWDHWFN 386

RESULT 8
US-10-313-203-10
; Sequence 10, Application US/10313203
; Publication No. US20040029134A1
; GENERAL INFORMATION:
; APPLICANT: Gu, Wei
; APPLICANT: Nikolaev, Anatoly
; TITLE OF INVENTION: P53-ASSOCIATED PARKIN-LIKE CYTOPLASMIC PROTEIN, AND RELATED COMPO
; TITLE OF INVENTION: AND METHODS
; FILE REFERENCE: 68106
; CURRENT APPLICATION NUMBER: US/10/313,203
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-313-203-10

Query Match      14.4%; Score 373; DB 15; Length 64;
Best Local Similarity 100.0%; Pred. No. 3e-25;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      314 RYQYGAEECVLQMGVYLCPRPCCAGLLPDPDRKVTCEGGANGLGGPAPFRECCEAYH 373
      |||
Db      1 RYQYGAEECVLQMGVYLCPRPCCAGLLPDPDRKVTCEGGANGLGGPAPFRECCEAYH 60
QY      374 EGEC 377
      |||
Db      61 EGEC 64

```

```

RESULT 9
US-10-313-203-17
; Sequence 17, Application US/10313203
; Publication No. US20040029134A1
; GENERAL INFORMATION:
; APPLICANT: Gu, Wei
; APPLICANT: Nikolaev, Anatoly
; TITLE OF INVENTION: P53-ASSOCIATED PARKIN-LIKE CYTOPLASMIC PROTEIN, AND RELATED COMPO
; FILE REFERENCE: 68106
; CURRENT APPLICATION NUMBER: US/10/313,203
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-313-203-17

Query Match      12.4%; Score 323; DB 15; Length 56;
Best Local Similarity 100.0%; Pred. No. 6.7e-21;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      238 CITCTVRSPLVAFQCN SRHVICLDCEFLYCVTRLNDROPVHDPOLGYSLEPCVAGC 293
      |||
Db      1 CITCTVRSPLVAFQCN SRHVICLDCEFLYCVTRLNDROPVHDPOLGYSLEPCVAGC 56

RESULT 10
US-09-864-761-36750
; Sequence 36750, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:

```

```
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aemica-X-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/532,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/508,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 36750
LENGTH: 46
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL035697.18
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.68
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.76
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.94
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.63
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.68
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.91
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.77
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.73
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.96
OTHER INFORMATION: EST_HUMAN HIT: T78558.1, EVALU8 2.00e+00
OTHER INFORMATION: SWISSPROT HIT: Q04833, EVALU8 9.10e-01
US-09-864-761-36750

Query Match 10.1%; Score 263; DB 9; Length 46;
Best Local Similarity 100.0%; Pred. No. 1e-15;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 312 YNRYYQYGAEECVLQWGVLCPRPGGAGLLPEPDQKVTCEGGNG 357
DB 1 YNRYYQYGAEECVLQWGVLCPRPGGAGLLPEPDQKVTCEGGNG 46
```

```
RESULT 11
US-10-437-963-199787
Sequence 199787, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Bardazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 199787
LENGTH: 525
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_95320C.1.pep
US-10-437-963-199787

Query Match 9.5%; Score 246; DB 16; Length 525;
Best Local Similarity 30.1%; Pred. No. 7.1e-13;
Matches 71; Conservative 26; Mismatches 93; Indels 46; Gaps 11;

QY 232 NSRNTCTCTCVRBPVLYFGNSRHYVCLDFHLYCVTRLNDROPVHPQGYSLPCV- 290
DB 127 NRRRLTCAICPVPFDTGGRSAGCSHYCVSCWGRGV-----RAVGDGARCLSRCPD 180
QY 291 ACCPNSLKEIHLHFRILGEEQYNYRQYGAEECVLQWGV-LCPRPCCGAGLLPEPDQK 349
DB 181 PSCPAAYVRELVDVADGDR-ERFGWFLRSLVSEBSAGMRWCPGCGC-----SRA 230
QY 350 VTCEGGNG-----LGGFAPCRBC-KEAYHEGCSAVFENSGTTTQAYRYDERAEO 400
DB 231 VEFVGGDGESESEVFCSCGHGLCWRCGEARPVSCKTV-----AKWYEKSSSES 281
QY 401 --ARWEASKEITIKTKPCRCCHVPEKNGCMHKCPQPCRLKEMCNCGEEN 454
DB 282 ETATWLLAH-----TRGCPKCRLLPIEKNLGCMHWTC-RPCLLHEFCWICLKPMH 329

RESULT 12
US-10-425-114-55407
Sequence 55407, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack B
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 55407
LENGTH: 604
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: LIB3051-106-A7_Flt1.pep
```

US-10-425-114-55407

Query Match 9.3%; Score 242.5; DB 15; Length 604;
Best Local Similarity 27.7%; Pred. No. 1.7e-12;
Matches 70; Conservative 31; Mismatches 97; Indels 55; Gaps 12;

QY 221 ETPALHLATNSNITCTTIVRSFVLVFOCNSRHVICTDCEHLYCTRLNDQ---- 276
DB 140 EKVFOQ---SNRELTGCI CFEMYPRAVESTACGHPYCSAGYFSTINDGPECVLV 196
QY 277 -FVHDPQLGYSLPCVACGPNLSIKELHFRILISEQYNRQYGAEECVLQMGV- LCPR 334
DB 197 LRCDEPS-----CGAAGQDMITLL-----ASDEDKQKTDRIYLSYISDNKKTWCPSA 245
QY 335 PGCGAGLLEPPDQKVTCEGNG-----LQGFAPFCREC-KEAYHEGECSAVFEASGTTT 388
DB 246 PGQ-----EYAVTFDAGSGNYDVSCISYFCMNCIEBAHPVDC-----GTVS 289
QY 389 QAVRVDERAABQARWEAASKETIKTKTFCPCRHVPEKNGGCMEMKCPQPCRLKWCYN 448
DB 290 KWLKNSAESENMMN-----ILANSKPCPKCRPIEKNOGCMHMTG-TPCKFEFCWL 341
QY 449 CGCEMNRVCMGDH 461
DB 342 CLGAMS-----DH 349

RESULT 13

US-10-080-608A-13
; Sequence 13, Application US/10080608A
; Publication No. US20030198956A1
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/10/080,608A
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 503
; TYPE: PRF
; ORGANISM: Drosophila melanogaster
US-10-080-608A-13

Query Match 9.3%; Score 240.5; DB 14; Length 503;
Best Local Similarity 27.5%; Pred. No. 2.1e-12;
Matches 72; Conservative 42; Mismatches 111; Indels 37; Gaps 13;

QY 209 FPKCGAHPISDKETPVALHLIATNSRNITCTTIVRSFVLVFOCNSRHVICTDCEHLYC 268
DB 105 FPKC-AHVINPNNATEAIKQKTSRSCCECEICFSQLEPDSMAGLECGHRFCMPCMHXL 163
QY 269 VTRLNDQPVHDPQLGYSLPCVA-GCPNLSIKELHFRILISEQYN-RYQYGAEECVLQ 326
DB 164 STK-----IVAEGIGQITSCAAGC-DILVDVTVALNLTDAVRVYQOOLITNSFV-E 215
QY 327 MGVVL--CPRPGCG-AGLLEPPDQKVTCEGNGLGGCFAPCRECKEAYHEG-ECSAVFE 382
DB 216 CNGLLRMWCSVDCTYAVVYPAEPRRHCK-----CGHVFCAFCGEMNHDVCKRWL-- 267
QY 383 ASGTTTQAVRVDERAABQARWEAASKETIKTKTFCPCRHVPEKNGGCMEMKCPQPCRL 442
DB 268 -----KKWIKKCCDDSETSNWIAAN-----TKCPCRCSTVTEKDGCCNHVVCNQNCK 315
QY 443 LEMCNCGCEMNRVCMGDH 464
DB 316 NEFCVCLGSWEP--HGSSWYN 335

RESULT 14

US-10-370-685-102

; Sequence 102, Application US/10370685
; Publication No. US20030215903A1
; GENERAL INFORMATION:
; APPLICANT: Hyman, Paul
; APPLICANT: Goldberg, Edward
; TITLE OF INVENTION: Nanostructures Containing PNA Joining and Functional Elements
; FILE REFERENCE: NANE.P-004
; CURRENT APPLICATION NUMBER: US/10/370,685
; PRIOR APPLICATION NUMBER: 10/080,608
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 102
; LENGTH: 503
; TYPE: PRF
; ORGANISM: Drosophila melanogaster
US-10-370-685-102

QY 209 FPKCGAHPISDKETPVALHLIATNSRNITCTTIVRSFVLVFOCNSRHVICTDCEHLYC 268
DB 105 FPKC-AHVINPNNATEAIKQKTSRSCCECEICFSQLEPDSMAGLECGHRFCMPCMHXL 163
QY 269 VTRLNDQPVHDPQLGYSLPCVA-GCPNLSIKELHFRILISEQYN-RYQYGAEECVLQ 326
DB 164 STK-----IVAEGIGQITSCAAGC-DILVDVTVALNLTDAVRVYQOOLITNSFV-E 215
QY 327 MGVVL--CPRPGCG-AGLLEPPDQKVTCEGNGLGGCFAPCRECKEAYHEG-ECSAVFE 382
DB 216 CNGLLRMWCSVDCTYAVVYPAEPRRHCK-----CGHVFCAFCGEMNHDVCKRWL-- 267
QY 383 ASGTTTQAVRVDERAABQARWEAASKETIKTKTFCPCRHVPEKNGGCMEMKCPQPCRL 442
DB 268 -----KKWIKKCCDDSETSNWIAAN-----TKCPCRCSTVTEKDGCCNHVVCNQNCK 315
QY 443 LEMCNCGCEMNRVCMGDH 464
DB 316 NEFCVCLGSWEP--HGSSWYN 335

Query Match 9.3%; Score 240.5; DB 15; Length 503;
Best Local Similarity 27.5%; Pred. No. 2.1e-12;
Matches 72; Conservative 42; Mismatches 111; Indels 37; Gaps 13;

QY 209 FPKCGAHPISDKETPVALHLIATNSRNITCTTIVRSFVLVFOCNSRHVICTDCEHLYC 268
DB 105 FPKC-AHVINPNNATEAIKQKTSRSCCECEICFSQLEPDSMAGLECGHRFCMPCMHXL 163
QY 269 VTRLNDQPVHDPQLGYSLPCVA-GCPNLSIKELHFRILISEQYN-RYQYGAEECVLQ 326
DB 164 STK-----IVAEGIGQITSCAAGC-DILVDVTVALNLTDAVRVYQOOLITNSFV-E 215
QY 327 MGVVL--CPRPGCG-AGLLEPPDQKVTCEGNGLGGCFAPCRECKEAYHEG-ECSAVFE 382
DB 216 CNGLLRMWCSVDCTYAVVYPAEPRRHCK-----CGHVFCAFCGEMNHDVCKRWL-- 267
QY 383 ASGTTTQAVRVDERAABQARWEAASKETIKTKTFCPCRHVPEKNGGCMEMKCPQPCRL 442
DB 268 -----KKWIKKCCDDSETSNWIAAN-----TKCPCRCSTVTEKDGCCNHVVCNQNCK 315
QY 443 LEMCNCGCEMNRVCMGDH 464
DB 316 NEFCVCLGSWEP--HGSSWYN 335

Query Match

9.0%; Score 233.5; DB 15; Length 585;

Best Local Similarity 26.5%; Pred. No. 1e-11;
Matches 67; Conservative 29; Mismatches 102; Indels 55; Gaps 12;

| | | | | | |
|----|-----|---------------|--|----|-----|
| QY | 221 | ETPVALLH | IATNSRNITCITCTDVRSPVLVFOGNSRHVICLDQFHLVCVTRIND---- | RQ | 276 |
| DB | 120 | EXPIVOH--- | PVARELTCGICFCENTPRARIEMASCGHPYRISCMWEGYISTSINDGPGCLM | | 176 |
| QY | 277 | FVHDPOLGYS | LPVAGCPNSLIKELHFRILGEEQYNRYOQYGAEECVL-QMGVLCPRP | | 335 |
| DB | 177 | LSPDPX----- | CGAATGRS-----SRFGSRSDEDQKQYARYLRLRSYIEDNKSKKCPAP | | 225 |
| QY | 336 | GCGAGLLPEPDQR | KVTCBSGN-----GIGCGPAFCREC-KEAYHEGECSAVFEEASGTTT | | 388 |
| DB | 226 | GC----- | EYAVTFDAGSTGNVDVSCLSYGFQMCNCTEBAHRPVDC-----GTVA | | 269 |
| QY | 389 | QAYRDERAAEQAR | WEAASKETIKKTKPCPRCHVPEKNGGCMHKCPOQCRLEWCMN | | 448 |
| DB | 270 | KWILKNSASESE | NNM-----ILANSKPCPKCKRPIEKNGCMHMTCT-TPPCKFEFCWL | | 321 |
| QY | 449 | CGCENNRVCMGDH | | | 461 |
| DB | 322 | CVGAWG----- | DH | | 329 |

Search completed: February 23, 2005, 14:33:24
Job time : 75.013 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using BW model

Run on: February 23, 2005, 13:34:50 ; Search time 19.3575 Seconds
(without alignments)
2311.294 Million cell updates/sec

Title: US-10-622-817-8

Perfect score: 2596
Sequence: 1 MIVFVRFNSSHGFPVEVDSD.....CNNCGCEMNRVCMDHMDV 465

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 490 | 18.9 | 357 | 2 T23460 | hypothetical prote |
| 2 | 220 | 8.5 | 543 | 2 H84724 | probable ARI-like |
| 3 | 219.5 | 8.5 | 1753 | 2 T00350 | hypothetical prote |
| 4 | 205 | 7.9 | 437 | 2 T16477 | hypothetical prote |
| 5 | 205 | 7.9 | 491 | 2 G87793 | protein C27A12.6 |
| 6 | 198.5 | 7.6 | 644 | 2 T02366 | hypothetical prote |
| 7 | 196.5 | 7.6 | 514 | 2 G84724 | probable ARI-like |
| 8 | 195 | 7.5 | 320 | 2 T45909 | hypothetical prote |
| 9 | 193 | 7.4 | 542 | 2 A84725 | similar to Atrialn |
| 10 | 190.5 | 7.3 | 436 | 2 H87793 | protein C27A12.8 |
| 11 | 188 | 7.2 | 1048 | 2 T31653 | hypothetical prote |
| 12 | 186.5 | 7.2 | 324 | 2 T16983 | hypothetical prote |
| 13 | 182.5 | 7.0 | 551 | 2 S38086 | hypothetical prote |
| 14 | 180.5 | 6.9 | 594 | 2 T04783 | hypothetical prote |
| 15 | 178.5 | 6.9 | 497 | 2 G87793 | protein C27A12.7 |
| 16 | 178.5 | 6.9 | 688 | 2 B86448 | hypothetical prote |
| 17 | 174.5 | 6.7 | 408 | 2 T47498 | hypothetical prote |
| 18 | 174 | 6.7 | 348 | 2 T47494 | hypothetical prote |
| 19 | 173 | 6.7 | 565 | 2 F84721 | probable RING zinc |
| 20 | 162.5 | 6.3 | 498 | 2 UC5983 | protein kinase C-1 |
| 21 | 159 | 6.1 | 518 | 2 E84536 | hypothetical prote |
| 22 | 158.5 | 6.1 | 451 | 2 T16481 | hypothetical prote |
| 23 | 157 | 6.0 | 488 | 2 T29562 | hypothetical prote |
| 24 | 155 | 6.0 | 532 | 2 T04748 | hypothetical prote |
| 25 | 152.5 | 5.9 | 816 | 2 T25555 | hypothetical prote |
| 26 | 149 | 5.7 | 464 | 2 S48329 | probable membrane |
| 27 | 147 | 5.7 | 1209 | 2 T52523 | hypothetical prote |
| 28 | 143.5 | 5.5 | 155 | 2 JH0227 | ubiquitin / riboso |
| 29 | 142.5 | 5.5 | 155 | 2 JH0226 | ubiquitin / riboso |

| | | | | | |
|----|-------|-----|-----|----------|---------------------|
| 30 | 141.5 | 5.5 | 468 | 2 A84601 | Mutator-like trans |
| 31 | 136 | 5.2 | 229 | 2 T47487 | hypothetical prote |
| 32 | 134.5 | 5.2 | 222 | 1 UQJN1 | ubiquitin precursor |
| 33 | 134 | 5.2 | 373 | 2 E84647 | hypothetical prote |
| 34 | 134 | 5.2 | 638 | 2 T02611 | hypothetical prote |
| 35 | 133 | 5.1 | 397 | 2 H84578 | probable RING zinc |
| 36 | 132.5 | 5.1 | 869 | 1 UC4858 | VLDL receptor prec |
| 37 | 132 | 5.1 | 80 | 2 T28305 | ORF MSV144 probabl |
| 38 | 132 | 5.1 | 150 | 2 T30390 | probable ubiquitin |
| 39 | 132 | 5.1 | 154 | 2 S55242 | polyubiquitin 2 - |
| 40 | 132 | 5.1 | 156 | 2 C86439 | protein T19E23.13 |
| 41 | 131 | 5.0 | 534 | 2 S34285 | polyubiquitin - to |
| 42 | 130.5 | 5.0 | 305 | 1 S29853 | polyubiquitin 4 - |
| 43 | 130.5 | 5.0 | 356 | 1 UQJTRC | polyubiquitin / ri |
| 44 | 130 | 5.0 | 155 | 2 S40240 | ubiquitin/ribosoma |
| 45 | 130 | 5.0 | 156 | 2 T52335 | ubiquitin extensio |

ALIGNMENTS

RESULT 1

T23460 hypothetical protein K08E3.7 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T23460
R/McMurray, A.
Submitted to the EMBL Data Library, November 1996
A/Reference number: Z19743
A/Accession: T23460
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-357 <NUL>
A/Cross-references: UNIPROT:Q9XUS3; EMBL:Z61568; PDB: CAB04599.1; GSPDB: GN00021; CESP: K08E3
A/Experimental source: clone K08E3
C/Genetics:
A/Gene: CESP:K08E3.7
A/Map position: 3
A/Intons: 23/3; 72/2; 218/1; 286/2; 311/2

| Query Match | 18.9% | Score 490 | DB 2 | Length 357 |
|-----------------------|-----------------|---|------------|------------|
| Best Local Similarity | 26.7% | Pred. No. 1.8e-29 | | |
| Matches 119 | Conservative 53 | Mismatches 154 | Indels 120 | Gaps 12 |
| QY | 22 | SIFOLKEVYAKKQGVADQLRYIFAKELRNDVTQNCDDQSIYHIYQRPARKQENW 81 | | |
| DB | 29 | NIEDLTQDVEKLEIPSDLELVFVCGKLSKSTIMRDLSTPATQIMLRKFNSHNENG 88 | | |
| QY | 82 | ATCGDDPRNAAGCEREPSQSLRVDSSTVLPDQSVGLAVILHTDSRKQSPAPAGRS 141 | | |
| DB | 89 | AT-----TAKITTS-----S 99 | | |
| QY | 142 | YNSFYVYKGPORVQPKLRVQSTCRQATLTLLQGPSQMDVLIPNMSGECQSPHC 201 | | |
| DB | 100 | ILGSFYVWKCN-CDVYKRLKRLRYCQKGSSTVLVSEBQMSDVLSKSRIPAVCECCT 158 | | |
| QY | 202 | PETSAEFPFKCGAHPSPDKETPVAAHLIATNSRNITCITDVRSVVLVFOQNSRIVICL 261 | | |
| DB | 159 | PGIFAEFKFKC-----LACNDPAALTHVRGNQMTECCVCDGKEKVIPIGLCN--HITQ 212 | | |
| QY | 262 | DCFHLYCVTRLNDROPVNDPQGLYSLPVAAGCNSLTIKELHPRIGSEQYNNYQYGA 321 | | |
| DB | 213 | FCFRV-----RVQDVNHHFIMGQSYSEYQKATE 243 | | |
| QY | 322 | ECV-LQMGVTLCPRPCGAGLLPEP--DQKVTCEGANGIGCFAPCRECKEAYHGECS 378 | | |
| DB | 244 | RLIAYDVKGVTCPPNNSCGQSFMEBPYDDGSGCP-----DCFFSFCRC--FRNCV 294 | | |
| QY | 379 | AVPEASGTTTQAVRVDERAABQARWBAASKETIKTKTKCPKCHVVEVAKGCGMAHKCFQ 438 | | |
| DB | 295 | COSEDDLTRT-----TIDATTRBCPKCHVATERNGCAHITHC-- 331 | | |

QY 439 POCRLWMCNCGCENNRVCMGDHMPD 464
DB 332 TSCGMDCWCFKCKTMEKECECOMDHEN 357

RESULT 2

H84724

Probable ARI-like RING zinc finger protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: H84724
R:Lin, X.; Keul, S.; Rounleay, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
eues, D.; Nierman, W.C.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.
Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: H84724
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-543 <STO>
A/Cross-references: UNIPROT:Q9SKC3; GB:AE002093; NID:94887759; PIDN:AA032295.1; GSPDB:GN
A/Gene: At2g31770
A/Map position: 2

Query Match
Best Local Similarity 25.9%; Score 220; DB 2; Length 543;
Matches 60; Conservative 33; Mismatches 83; Indels 56; Gaps 11;

QY 235 NITCITCTDVRSFVLVFCQNSRHVLCDFHLYCVTRIND-----RQFVNDPQLGYSLPC 289
DB 125 NIQCGICFESYRREIARVSCGHPYCKTCMAGITTKIEDPGCLRYKCEPPS-----C 178
QY 290 VAGCPNSLIXELHHPILIGEOYNY--QQTGAEECVLQMGVLCPPRGAGALL--PE 344
DB 179 SAAVGMDIEDYETKV--NEKTSRYILRSYVEGKKIKW----CPSPGGVAVNEGGSF 232
QY 345 PQRKYTEGGNGLCGFAFCRECKEAYHEG--ECSAVFBAISGTTQAYRDERAABOARM 403
DB 233 SSSYDVSLC-----CSYRFCMCSBDHSHVDDTY-----SKW 266
QY 404 -----EASKEIKTKTKRCPCHVVEKNGGCMHMKCPQPCRLEWMCNC 449
DB 267 IFKQDSEENKMWLANSKPCPECKRPIEHKNDGNHMTCSAP--CGHEFCWIC 317

RESULT 3

T00350

hypothetical protein KIA0708 - human (fragment)
C:Species: Homo sapiens (man)
C/Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C/Accession: T00350
R:Ichikawa, K.; Nagase, T.; Shiyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A>Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A/Reference number: Z14142; MUID:98403880; PMID:9734811
A/Accession: T00350
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1753 <ISH>
A/Cross-references: UNIPROT:O81WT3; EMBL:AB014608; NID:93327229; PIDN:BA031683.1; PID:G3
A/Experimental source: brain
A/Genetics:
A/Note: KIA0708

Query Match
Best Local Similarity 8.5%; Score 219.5; DB 2; Length 1753;
Matches 125; Conservative 57; Mismatches 181; Indels 221; Gaps 32;

QY 1 MIVVPRNRSISGFPEVY--DSDTS-----IFOLK 27
DB 1012 MWLLKRNQTEEVSVETLKDSDLSPELLQLMLVPLTSGNGPLTLHGGDFPHGGVLRILH 1071

QY 28 EVVAKROG-----VPADQLRVIFAGKELRNDTV-----QNCDDLPQOSI 66
DB 1072 ERGPQRSEALMLIPQAYLVNVEDEGRITLQKRNLISCLVRLKAKHGEKXHLHIDQ--L 1129

QY 67 VHIYORPWRKQENMNAAGDPRNAAGCEREPQSLTRVDSVSVLPFGDSVGLAVIHLTD 126
DB 1130 VCLVLEAMOKGPNPPTLIG--HTVAGV-----ACTSIDVLSCTL--HLGGQYV----- 1175

QY 127 SRKDSPP-----AGSPAGSINSTRVYCKPCQRPVOPGKLRVQSTCRQATLTITGCP 180
DB 1176 KRDRDRQILMVAAPRPMGRCQADVPFCGQSETSKRPEAV-----ATLASQLP 1228

QY 181 SCMDVLIENRMSGECSPHCPGTSAEFFFK--CGAIPTSDKTTPVALHILATNS---- 233
DB 1229 -----AGRTMSPO-----EVEGLMQYTRQVETLNLBPVVAQHILASHWGAE 1272

QY 234 -----NITCITCTDVRSFV-----LVFQNSRHVI 259
DB 1273 QLLQSYSEDPPEPLLAAGLCVHQAOAVPRBDHPCV--SPLGCDLPLSLC--CMHYC 1328

QY 260 CLDCHLYCVTRINDRQFVNDPQLGYSLPC--VAGCP--NSLIKELHFRILIGEOYNR 314
DB 1329 CKSCWNEVILTRT-----EQNLVINTCPRIADCPAPGATIRAI--VSSPEVSK 1377

QY 315 YQO-----YGAEECVLQMGVLCPP--GCGAGLLPBDQRKYTEGGNGILG-----CGFA 363
DB 1378 YEKALLRQY--VESC--SNLTWCTNPGCD-----RLICRQGLGCGITCKCKGWA 1423

QY 364 FCRBCK--EAVHEGCGAVFEASGTTTQAYRDERAABOARM-----EASKE 409
DB 1424 SCFNCSEPEAHYPAAC-----GHNSQVVDGQYGYDGKSVYDAOSKH 1463

QY 410 TIKKTRPCRCVPEYKNGGCMHMKCPQPCRLEWMCNCGEW 453
DB 1464 LAKLISRCRSCQAPLEKNEGCLHMTG--AKCNHGFWRCLIKSW 1505

RESULT 4

T16477

hypothetical protein F56D2.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C/Accession: T16477
R:Du, Z.

submitted to the EMBL Data Library, August 1994
A/Description: The sequence of C. elegans cosmid F56D2.
A/Reference number: Z18519
A/Accession: T16477

A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA

A/Residues: 1-437 <DUZ>
A/Cross-references: UNIPROT:Q20871; EMBL:U13644; NID:9532100; PID:91945502; PIDN:AA85268-

C/Genetics:
A/Experimental source: strain Bristol N2; clone F56D2
A/Gene: CESP:F56D2.5
A/Map position: 3

A/Note: 41/3; 134/3; 185/3; 223/3; 254/2; 287/2; 414/3

Query Match
Best Local Similarity 7.9%; Score 205; DB 2; Length 437;
Matches 104; Conservative 66; Mismatches 169; Indels 176; Gaps 28;

QY 18 DSDTSIFOLKEVAVAKQGVADQLRVIFAGKEL--RNDWTQNCDDLDQOSTVHIVGRPMR 75
DB 3 DRDLQIYEL-----EALSVAREKKLAKSSWSDSKNAEI--QGIIEV----- 42

QY 76 KQGEWNAATGDDPRNAAGCEREPQSLTRVDSVSVLPFGDSVGLAVIHLTDSKRDSPAG 135
DB 43 -----GFDNLVDPYTTITGSDSGDFHLPLDILP-----PIRLKHLNNDVPTYS 88

QY 136 SPAGRSIYNSFYVCKGFCQVRGKLRVOCSTCRQATLTLTGQSPCGMDVLIENRMSG 195
DB 136 SPAGRSIYNSFYVCKGFCQVRGKLRVOCSTCRQATLTLTGQSPCGMDVLIENRMSG 195

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:25:39 / Search time 102.258 Seconds
(without alignment)
1758.725 Million cell updates/sec

Title: US-10-622-817-9
Perfect score: 2596
Sequence: 1 MIVFRFNSHGFPVEVDSI.....CWNCCENRVCMDGHFVDV 465

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues
Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : A_Geneseq16Dec04:*

- 1: geneseqp19808:*
- 2: geneseqp19908:*
- 3: geneseqp20008:*
- 4: geneseqp20018:*
- 5: geneseqp20028:*
- 6: geneseqp20038:*
- 7: geneseqp20038:*
- 8: geneseqp20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|--------------------|
| 1 | 2596 | 100.0 | 465 | 2 | AAV32501 Human par |
| 2 | 2596 | 100.0 | 465 | 6 | AAE30800 Human par |
| 3 | 2596 | 100.0 | 465 | 6 | ABO07156 Human p53 |
| 4 | 2584 | 99.5 | 465 | 6 | AAE30801 Human par |
| 5 | 2405 | 92.6 | 437 | 2 | AAV32502 Human par |
| 6 | 2405 | 92.6 | 437 | 6 | ABO07157 Human p53 |
| 7 | 2170.5 | 83.6 | 464 | 4 | AAE67517 Amino aci |
| 8 | 2165.5 | 83.4 | 464 | 4 | AAE67531 Amino aci |
| 9 | 2165.5 | 83.4 | 464 | 4 | AAE67532 Amino aci |
| 10 | 2087.5 | 80.4 | 451 | 4 | AAE67533 Amino aci |
| 11 | 1696.5 | 65.4 | 316 | 6 | ABO07158 Human p53 |
| 12 | 1679.5 | 64.7 | 344 | 4 | AAE67521 Amino aci |
| 13 | 1234.5 | 47.6 | 296 | 4 | AAE67526 Amino aci |
| 14 | 1058.5 | 40.8 | 468 | 4 | ABE65114 Drosophi |
| 15 | 1002 | 38.6 | 250 | 4 | AAE67519 Amino aci |
| 16 | 997 | 38.4 | 262 | 4 | AAE67518 Amino aci |
| 17 | 870 | 33.5 | 156 | 4 | AAE67525 Amino aci |
| 18 | 671.5 | 25.9 | 183 | 4 | AAE67522 Amino aci |
| 19 | 671.5 | 25.9 | 194 | 4 | AAE67524 Amino aci |
| 20 | 544.5 | 21.0 | 386 | 4 | AAE6951 C.elegan |
| 21 | 506 | 19.5 | 153 | 4 | AAE67523 Amino aci |
| 22 | 335 | 12.9 | 77 | 4 | AAE67529 Amino aci |
| 23 | 263 | 10.1 | 46 | 4 | AAE17055 Peptide # |
| 24 | 263 | 10.1 | 46 | 4 | ABE36052 Peptide # |
| 25 | 263 | 10.1 | 46 | 4 | AAE29548 Peptide # |

| | | | | | |
|----|-------|------|-----|---|---------------------|
| 26 | 263 | 10.1 | 46 | 4 | ABE30873 Peptide # |
| 27 | 263 | 10.1 | 46 | 4 | ABE21452 Protein # |
| 28 | 263 | 10.1 | 46 | 4 | AAE69222 Human bon |
| 29 | 263 | 10.1 | 46 | 4 | AAE56839 Human bra |
| 30 | 263 | 10.1 | 46 | 4 | ABE50898 Human liv |
| 31 | 263 | 10.1 | 46 | 5 | ABE38830 Human pep |
| 32 | 252 | 9.7 | 106 | 8 | ADN96489 Modified |
| 33 | 251 | 9.7 | 63 | 4 | AAE67522 Amino aci |
| 34 | 251 | 9.7 | 105 | 4 | AAE67520 Amino aci |
| 35 | 251 | 9.7 | 105 | 8 | ADN96487 Modified |
| 36 | 240.5 | 9.3 | 503 | 4 | ABE61708 Drosophi |
| 37 | 240.5 | 9.3 | 503 | 7 | ADN67082 Atrialne-1 |
| 38 | 240 | 9.2 | 511 | 4 | ABE63665 Drosophi |
| 39 | 239 | 9.2 | 520 | 8 | ADP22562 Sea-squit |
| 40 | 236.5 | 9.1 | 445 | 3 | AAE21034 Human nuc |
| 41 | 236.5 | 9.1 | 445 | 8 | ADN96154 T cell ac |
| 42 | 236.5 | 9.1 | 445 | 8 | ADN96026 T cell ac |
| 43 | 236.5 | 9.1 | 557 | 3 | AAE98059 Human Rn |
| 44 | 236.5 | 9.1 | 557 | 8 | ADN96028 T cell ac |
| 45 | 236 | 9.1 | 53 | 4 | ABE67528 Amino aci |

ALIGNMENTS

| | | |
|-----------------------|---|----------------------------|
| RESULT 1 | AAV32501 | standard; protein; 465 AA. |
| ID | AAV32501 | standard; protein; 465 AA. |
| XX | AAV32501; | |
| AC | AAV32501; | |
| XX | AAV32501; | |
| DT | 21-OCT-1999 | (first entry) |
| XX | 21-OCT-1999 | (first entry) |
| DE | Human parkin gene variant protein. | |
| XX | Human parkin gene variant protein. | |
| KW | Parkin's disease related gene; parkin gene; variant; gene therapy. | |
| XX | Parkin's disease related gene; parkin gene; variant; gene therapy. | |
| OS | Homo sapiens. | |
| XX | Homo sapiens. | |
| PN | WO9940191-A1. | |
| XX | WO9940191-A1. | |
| PD | 12-AUG-1999. | |
| XX | 12-AUG-1999. | |
| PF | 09-FEB-1999; | 99WO-JP000545. |
| XX | 09-FEB-1999; | 99WO-JP000545. |
| PR | 09-FEB-1998; | 98JP-00027531. |
| XX | 09-FEB-1998; | 98JP-00027531. |
| PA | (SHIM/) SHIMIZU N. | |
| XX | (SHIM/) SHIMIZU N. | |
| PI | (MIZU/) MIZUNO Y. | |
| XX | (MIZU/) MIZUNO Y. | |
| XX | Shimizu N, Mizuno Y; | |
| DR | WPI; 1999-494295/41. | |
| XX | WPI; 1999-494295/41. | |
| DR | N-PSDB; AAE99923. | |
| XX | N-PSDB; AAE99923. | |
| PT | Gene implicated in the pathology of Parkinson's disease, used for | |
| XX | Gene implicated in the pathology of Parkinson's disease, used for | |
| PT | treatment of the disease. | |
| XX | treatment of the disease. | |
| PS | Claim 1; Page 83-88; 14pp; English. | |
| XX | Claim 1; Page 83-88; 14pp; English. | |
| CC | This sequence is encoded by a gene of the invention, and is implicated in | |
| XX | This sequence is encoded by a gene of the invention, and is implicated in | |
| CC | the pathology of Parkinson's disease. This sequence is a variant of the | |
| XX | the pathology of Parkinson's disease. This sequence is a variant of the | |
| CC | parkin gene found in parkinson's disease patients. The sequences may be | |
| XX | parkin gene found in parkinson's disease patients. The sequences may be | |
| CC | used for the diagnosis, treatment (including gene therapy) and | |
| XX | used for the diagnosis, treatment (including gene therapy) and | |
| CC | investigation of Parkinson's disease | |
| XX | investigation of Parkinson's disease | |
| SQ | Sequence 465 AA; | |
| XX | Sequence 465 AA; | |
| Query Match | 100.0%; Score 2596; DB 2; Length 465; | |
| Best Local Similarity | 100.0%; Pred. No. 2.3e-219; | |
| Matches | 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | |
| QY | 1 MIVFRFNSHGFPVEVDSITIFQLKEVVAKRGVAPADQLRVIFAGKELRMDWTQVNC | 60 |

```

Db      1 MIVFVRNSSHGFPEVSDTSTIFQLKEVAKRGQVPAQRLVIFAGKELRNDMTVQNC
QY      61 LDQOSIVHIVORPMKQGEENATGDDPRNAAGCCEREPOSILTRVLDSSSVLPQDSVGLA
Db      61 LDQOSIVHIVORPMKQGEENATGDDPRNAAGCCEREPOSILTRVLDSSSVLPQDSVGLA
QY      121 VILHTRKSDSPAPGSPAGRSIYNSFYVYCKGCPQVOPGKLRVQCSCTCQATLTLTQGP
Db      121 VILHTRKSDSPAPGSPAGRSIYNSFYVYCKGCPQVOPGKLRVQCSCTCQATLTLTQGP
QY      181 SCMDVLIIPNRMSECCSPHCPGTSAPFFKCGAHPTSDKETPVALLIATNSRNTTCT
Db      181 SCMDVLIIPNRMSECCSPHCPGTSAPFFKCGAHPTSDKETPVALLIATNSRNTTCT
QY      241 CTDVRSPLVLFQCNRSRHVILCLDFHLYCVTRLNDROFVHDPOGSLPCVACCPNSLKE
Db      241 CTDVRSPLVLFQCNRSRHVILCLDFHLYCVTRLNDROFVHDPOGSLPCVACCPNSLKE
QY      301 LHHFRILGEBQVNRVQOYGAEECVLQMGVLCPRPGCAGLLPEPDQRKVTCEGNGLGC
Db      301 LHHFRILGEBQVNRVQOYGAEECVLQMGVLCPRPGCAGLLPEPDQRKVTCEGNGLGC
QY      361 GFAPFCRECKEAYHEGECSAVFEASGTTTQAYRVDERRAABQARWEAASKETTKTTKPCR
Db      361 GFAPFCRECKEAYHEGECSAVFEASGTTTQAYRVDERRAABQARWEAASKETTKTTKPCR
QY      421 CHVPEKNGGCMHMKCPQPCRLIEWCNCCEMNRCVCMGDHMFV
Db      421 CHVPEKNGGCMHMKCPQPCRLIEWCNCCEMNRCVCMGDHMFV

```

RESULT 2

AAE30800 ID AAE30800 standard; protein; 465 AA.

AAE30800;

24-FEB-2003 (first entry)

Human Parkin protein.

Human; Parkin protein; neurological disorder; apoptosis; gene therapy; ischaemic stroke; Parkinson's disease; Alzheimer's disease; nocturnal; transgenic; cerebrotectonic; neuroprotective; neurotransplantation.

Homo sapiens.

Key Location/Qualifiers

Cleavage-site 126

MO300279459-A2.

10-OCT-2002.

02-APR-2002; 2002MO-DK000221.

29-MAR-2001; 2001DK-00000525.

03-APR-2001; 2001US-0281286P.

(NSG-) NSGENE AS.

Jensen PH;

WPI; 2003-046812/04.

N-PSDB; AAD47679.

New isolated nucleic acid sequence encoding a Parkin polypeptide, useful for treating, preventing or diagnosing neurological disorders, e.g. Parkinson's disease, Alzheimer's disease or ischemic stroke, and in screening assays.

Claim 10; Page 69; 71pp; English.

XX The invention relates to Parkin protein and its corresponding nucleic acid sequence. The nucleic acid sequence is useful for altering the proteolytic processing of Parkin at its potential cleavage site at Asp 126. The invention is used in manufacturing or testing a pharmaceutical composition for treating and/or preventing a neurological disorder, e.g. Alzheimer's disease or ischemic stroke. It also used for detecting the occurrence of proteolytic processing of Parkin at Asp 126 in a sample, in monitoring a potential disposition for a neurodegenerative disease, and for treating, preventing and/or diagnosing Parkinson's disease or other neurodegenerative disorders. The viral vector is used for transforming neuronal cells in vivo or ex vivo. The invention is used for transforming neurotransplantation into the CNS of a mammal. It may be used in screening assays to identify compounds that increase or decrease apoptosis. It is also used in gene therapy. The present sequence is human Parkin protein.

Sequence 465 AA;

Query Match 100.0%; Score 2596; DB 6; Length 465; Best Local Similarity 100.0%; Pred. No. 2.3e-219; Mismatches 0; Gaps 0; Matches 465; Conservative 0; Indels 0;

```

QY      1 MIVFVRNSSHGFPEVSDTSTIFQLKEVAKRGQVPAQRLVIFAGKELRNDMTVQNC
Db      1 MIVFVRNSSHGFPEVSDTSTIFQLKEVAKRGQVPAQRLVIFAGKELRNDMTVQNC
QY      61 LDQOSIVHIVORPMKQGEENATGDDPRNAAGCCEREPOSILTRVLDSSSVLPQDSVGLA
Db      61 LDQOSIVHIVORPMKQGEENATGDDPRNAAGCCEREPOSILTRVLDSSSVLPQDSVGLA
QY      121 VILHTRKSDSPAPGSPAGRSIYNSFYVYCKGCPQVOPGKLRVQCSCTCQATLTLTQGP
Db      121 VILHTRKSDSPAPGSPAGRSIYNSFYVYCKGCPQVOPGKLRVQCSCTCQATLTLTQGP
QY      181 SCMDVLIIPNRMSECCSPHCPGTSAPFFKCGAHPTSDKETPVALLIATNSRNTTCT
Db      181 SCMDVLIIPNRMSECCSPHCPGTSAPFFKCGAHPTSDKETPVALLIATNSRNTTCT
QY      241 CTDVRSPLVLFQCNRSRHVILCLDFHLYCVTRLNDROFVHDPOGSLPCVACCPNSLKE
Db      241 CTDVRSPLVLFQCNRSRHVILCLDFHLYCVTRLNDROFVHDPOGSLPCVACCPNSLKE
QY      301 LHHFRILGEBQVNRVQOYGAEECVLQMGVLCPRPGCAGLLPEPDQRKVTCEGNGLGC
Db      301 LHHFRILGEBQVNRVQOYGAEECVLQMGVLCPRPGCAGLLPEPDQRKVTCEGNGLGC
QY      361 GFAPFCRECKEAYHEGECSAVFEASGTTTQAYRVDERRAABQARWEAASKETTKTTKPCR
Db      361 GFAPFCRECKEAYHEGECSAVFEASGTTTQAYRVDERRAABQARWEAASKETTKTTKPCR
QY      421 CHVPEKNGGCMHMKCPQPCRLIEWCNCCEMNRCVCMGDHMFV
Db      421 CHVPEKNGGCMHMKCPQPCRLIEWCNCCEMNRCVCMGDHMFV

```

RESULT 3

AB007156 ID AB007156 standard; protein; 465 AA.

AB007156;

13-AUG-2003 (first entry)

Human p53 modifying protein, SEQ ID 116.

Human; p53 modifier; cytostatic; cancer; cytostatic; antiangiogenic; antiapoptotic; p53 pathway; breast cancer; colon cancer; kidney cancer; lung cancer; ovarian cancer; angiogenesis; cell cycle; apoptotic disorder; cell proliferation disorder.

Homo sapiens.

| | |
|----|---|
| Pt | MO200299122-AL. |
| Pd | 12-DEC-2002. |
| Pf | 03-JUN-2002; 2002WO-US017382. |
| Pg | 05--JUN-2001; 2001US-0296076P. |
| Pr | 10-OCT-2001; 2001US-0328605P. |
| Px | 15-FEB-2002; 2002US-0357253P. |
| Xx | (EXEL-) EXELIXIS INC. |
| Pa | Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP, |
| Pi | WPI; 2003-156859/15. |
| Dx | N-P5DB; ACDI3332. |
| Dr | |
| Pt | Identifying modulators of the p53 pathway for use in treating apoptotic |
| Pt | or cell proliferation disorders, comprises screening for agents that |
| Pt | modulate activity of a human ortholog of genes that modify the p53 |
| Pt | pathway in Drosophila. |
| Pt | |
| Pt | Example 2; Page 399-401, 678bp; English. |
| Pt | |
| Pt | The invention relates to identifying (M1) a candidate p53 pathway |
| Pt | modulating agent, by contacting an assay system comprising a purified HM |
| Pt | polypeptide (human orthologue of genes that modify the p53 pathway in |
| Pt | Drosophila) or nucleic acid with a test agent under conditions, where but |
| Pt | for the presence of the test agent, the system provides a reference |
| Pt | activity, and detecting a test agent-biased activity of the assay system. |
| Pt | Also included are modulating (M2) a p53 pathway of a cell (comprising |
| Pt | contacting a cell defective in p53 function with a candidate modulator |
| Pt | that specifically binds to a HM polypeptide comprising an HM amino acid |
| Pt | sequence, where p53 function is restored), modulating (M3) a p53 pathway |
| Pt | in a mammalian cell (comprising contacting the cell with an agent that |
| Pt | specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4) |
| Pt | a disease in a patient (comprising: (a) obtaining a biological sample |
| Pt | from the patient; (b) contacting the sample with a probe for HM |
| Pt | expression; (c) comparing the results with a control; and (d) determining |
| Pt | whether the comparison indicates a likelihood disease). (M1) is useful |
| Pt | for identifying modulators of the p53 pathway. A probe for HM expression |
| Pt | is useful for diagnosing breast, colon, kidney, lung and ovarian cancer, |
| Pt | in a patient, where the cancer has greater than 25 % expression level. |
| Pt | Modulators identified by (M1) are useful in a variety of diagnostic and |
| Pt | therapeutic applications, where disease or disorder prognosis is related |
| Pt | to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell |
| Pt | proliferation disorders (e.g. cancer). Another two new methods (M2 and |
| Pt | M3) are useful for modulating the p53 pathway of a cell, thus restoring |
| Pt | the p53 function of the cell, so that the cell undergoes normal |
| Pt | proliferation or progression through the cell cycle. (M2) and (M3) are |
| Pt | also useful for treating defects in the p53 pathway such as angiogenic, |
| Pt | apoptotic or cell proliferation disorder. The present sequence |
| Pt | represents a human p53 pathway modifying protein |
| Pt | |
| Pt | Sequence 465 AA; |
| Pt | |
| Pt | Query Match 100.0%; Score 2596; DB 6; Length 465; |
| Pt | Best Local Similarity 100.0%; Pred. No. 2.3e-219; |
| Pt | Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0 |
| Pt | |
| Pt | 1 MIVFVFNSSHQFPVEVDSDTISIPOLKEVNARKQGVPAQLRVIFAGKELRNDMTVQNCD 60 |
| Pt | 1 MIVFVRFNSSHGFPEYVDSDTISIPOLKEVNARKQGVPAQLRVIFAGKELRNDMTVQNCD 60 |
| Pt | |
| Pt | 61 LDQGSIVHIVQRPMKQGEMNATGDDPRNNAAGCGEREPQSILTRVDLSSSVLPGRSVGLA 120 |
| Pt | 61 LDQGSIVHIVQRPMKQGEMNATGDDPRNNAAGCGEREPQSILTRVDLSSSVLPGRSVGLA 120 |
| Pt | |
| Pt | 61 LDQGSIVHIVQRPMKQGEMNATGDDPRNNAAGCGEREPQSILTRVDLSSSVLPGRSVGLA 120 |
| Pt | 61 LDQGSIVHIVQRPMKQGEMNATGDDPRNNAAGCGEREPQSILTRVDLSSSVLPGRSVGLA 120 |
| Pt | |
| Pt | 121 VILHTDSRKDSPAPASPGASRTYNFFYYCKPCQCRVOPGSKLRVOCFCRCROATLTLTQGP 180 |
| Pt | 121 VILHTDSRKDSPAPASPGASRTYNFFYYCKPCQCRVOPGSKLRVOCFCRCROATLTLTQGP 180 |
| Pt | |
| Pt | 181 SCWDVLVILPNRMSCGCQSPHCSTGAEEFFKCQAHPTSDKETPVALHLIATNSRNITCIT 240 |
| Pt | 181 SCWDVLVILPNRMSCGCQSPHCSTGAEEFFKCQAHPTSDKETPVALHLIATNSRNITCIT 240 |

| | | | |
|----------|---|--|-----|
| Db | 181 | SCDDVILIPKRNAGEEQSHPCRTSAEPFKGALPTSDKETPVALLIANSNTICIT | 240 |
| Qy | 241 | CTDVRSVPLVFGQNSRHVI CLDQFHLVCTYRLNDRQFVHDPQLGYLPCVAGCPNSLIKE | 300 |
| Db | 241 | CTDVRSVPLVFGQNSRHVI CLDQFHLVCTYRLNDRQFVHDPQLGYLPCVAGCPNSLIKE | 300 |
| Qy | 301 | LHFRRLIGEEQYNNRYQYGAEECVYLQMGVLCPRPGCGAGLLPEPDQKRYTCGGNGLGC | 360 |
| Db | 301 | LHFRRLIGEEQYNNRYQYGAEECVYLQMGVLCPRPGCGAGLLPEPDQKRYTCGGNGLGC | 360 |
| Qy | 361 | GPAFCRECKEAHYEGSCSAVFPAESGTTTQAYRVDEBAAEQARWEASKETIKTTTKPCR | 420 |
| Db | 361 | GPAFCRECKEAHYEGSCSAVFPAESGTTTQAYRVDEBAAEQARWEASKETIKTTTKPCR | 420 |
| Qy | 421 | CHYPVEKNQSCMHMKCPQPCRLBMCNMGCEMNVCNGDHFVY 465 | |
| Db | 421 | CHYPVEKNQSCMHMKCPQPCRLBMCNMGCEMNVCNGDHFVY 465 | |
| RESULT 4 | | | |
| AAE30801 | | | |
| ID | AAE30801 | standard; protein; 465 AA. | |
| AC | AAE30801; | | |
| XX | | | |
| DT | 24-FEB-2003 | (first entry) | |
| XX | | | |
| DE | | Human Parkin D126 mutant protein. | |
| XX | | | |
| KW | | Human; Parkin protein; neurological disorder; apoptosis; gene therapy; | |
| KW | | Ischaemic stroke; Parkinson's disease; Alzheimer's disease; neurologic; | |
| KW | | transgenic; cerebroprotective; neuroprotective; neurotransplantation; | |
| KW | | mutant; muten. | |
| XX | | | |
| OS | | Homo sapiens. | |
| OS | | Synthetic. | |
| XX | | | |
| XX | | | |
| Key | | Location/Qualifiers | |
| FT | Misc-difference 126 | | |
| FT | /note= "Wild-type Asp is replaced with Glu" | | |
| FT | Cleavage-site 126 | | |
| FT | Misc-difference 223 | | |
| FT | /note= "Encoded by CCA" | | |
| XX | | | |
| PN | W0200279459-A2. | | |
| XX | | | |
| PD | 10-OCT-2002. | | |
| XX | | | |
| PF | 02-APR-2002; 2002WO-DK000221. | | |
| XX | | | |
| PR | 29-MAR-2001; 2001DK-00000525. | | |
| PR | 03-APR-2001; 2001US-0281286P. | | |
| XX | | | |
| PA | (NSGE-) NSGENE AS. | | |
| XX | | | |
| PI | Jensen PH; | | |
| XX | | | |
| DR | WPI; 2003-046812/04. | | |
| DR | N-PSDB; AAD47680. | | |
| XX | | | |
| PT | | New isolated nucleic acid sequence encoding a Parkin polypeptide, useful | |
| PT | | for treating, preventing or diagnosing neurological disorders, e.g. | |
| PT | | Parkinson's disease, Alzheimer's disease or ischemic stroke, and in | |
| PT | | screening assays. | |
| XX | | | |
| PS | Claim 10; Page 71; 71pp; English. | | |
| XX | | | |
| CC | | The invention relates to Parkin protein and its corresponding nucleic | |
| CC | | acid sequence. The nucleic acid sequence is useful for altering the | |
| CC | | proteolytic processing of Parkin at its potential cleavage site at Asp | |
| CC | | 126. The invention is used in manufacturing or testing a pharmaceutical | |
| CC | | composition for treating and/or preventing a neurological disorder, e.g. | |

CC Alzheimer's disease or ischaemic stroke. It also used for detecting the
CC occurrence of proteolytic processing of parkin at Asp 126 in a sample, in
CC monitoring a potential disposition for a neurodegenerative disease, and
CC for treating, preventing and/or diagnosing Parkinson's disease or other
CC neurodegenerative disorders. The viral vector is used for transforming
CC neuronal cells in vivo or ex vivo. The invention is useful for
CC neurotransplantation into the CNS of a mammal. It may be used in
CC screening assays to identify compounds that increase or decrease
CC apoptosis. It is also used in gene therapy. The present sequence is human
CC parkin mutant protein
XX

Seq Sequence 465 AA;

Query Match 99.5%; Score 2584; DB 6; Length 465;
Best Local Similarity 99.6%; Pred. No. 2.6e-218;

Matches 463; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MIVFVFNSSHGFPVEVDSDTSLFOLKEVNAKRGVPADQLRVIFAGKELRNDMTVQNC 60
DB 1 MIVFVFNSSHGFPVEVDSDTSLFOLKEVNAKRGVPADQLRVIFAGKELRNDMTVQNC 60
QY 61 LDQOSIVHIVORPMRKGEMNATGDDPRNNAAGCEREPOSTLRVDSSTVLPGDSVGLA 120
DB 61 LDQOSIVHIVORPMRKGEMNATGDDPRNNAAGCEREPOSTLRVDSSTVLPGDSVGLA 120
QY 121 VILHTDSRSDSPAGSPAGRSITNSFYCYCKGPCQORVOPGKLRVQSTCRQATLTLTQGP 180
DB 121 VILHTDSRSDSPAGSPAGRSITNSFYCYCKGPCQORVOPGKLRVQSTCRQATLTLTQGP 180
QY 181 SCMDVLLIPNRMSEGCSPHCPGTSABEFFPKCAHPTSDKETPVVALHLIATNSRNTTCT 240
DB 181 SCMDVLLIPNRMSEGCSPHCPGTSABEFFPKCAHPTSDKETPVVALHLIATNSRNTTCT 240
QY 241 CTDVRSPLVLFQCNRSRHVITCLDCFHLVCYTRLNDRFVHDPOLGYSILPCVAGCPNLSLKE 300
DB 241 CTDVRSPLVLFQCNRSRHVITCLDCFHLVCYTRLNDRFVHDPOLGYSILPCVAGCPNLSLKE 300
QY 301 LHHFRILGEBQYNNRYOYGABECVLOMGVLCPRPGCAGLLPEPDQRKVTCEGNGLGC 360
DB 301 LHHFRILGEBQYNNRYOYGABECVLOMGVLCPRPGCAGLLPEPDQRKVTCEGNGLGC 360
QY 361 GFACRECKEAYHGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKTTKPCPR 420
DB 361 GFACRECKEAYHGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKTTKPCPR 420
QY 421 CHVPEKNGGCMHMKCPQPCRLIEWCNCGEEMNRVCMGDHMFV 465
DB 421 CHVPEKNGGCMHMKCPQPCRLIEWCNCGEEMNRVCMGDHMFV 465

RESULT 5
AAV32502
ID AAV32502 standard; protein; 437 AA.

XX AAV32502;

DT 21-OCT-1999 (first entry)

XX Human parkin gene variant protein.

XX Parkinson's disease related gene; parkin gene; variant; gene therapy.

OS Homo sapiens.

XX WO940191-A1.

XX 12-AUG-1999.

XX 09-FEB-1999; 99WO-JP000545.

XX 09-FEB-1998; 98JP-00027531.

XX (SHIMU) SHIMIZU N.

PA (MIZU) MIZUNO Y.

XX Shimizu N, Mizuno Y;

XX WPI: 1999-494295/41.

XX N-PSDB; AAX99924.

XX Gene implicated in the pathology of Parkinson's disease, used for
XX treatment of the disease.

XX Claim 1; Page 89-94; 114pp; English.

CC This sequence is encoded by a gene of the invention, and is implicated in
CC the pathology of Parkinson's disease. This sequence is a variant of the
CC parkin gene found in parkinson's disease patients. The sequences may be
CC used for the diagnosis, treatment (including gene therapy) and
CC investigation of Parkinson's disease
XX

Seq Sequence 437 AA;

Query Match 92.6%; Score 2405; DB 2; Length 437;
Best Local Similarity 94.0%; Pred. No. 1.3e-202;

Matches 437; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

QY 1 MIVFVFNSSHGFPVEVDSDTSLFOLKEVNAKRGVPADQLRVIFAGKELRNDMTVQNC 60
DB 1 MIVFVFNSSHGFPVEVDSDTSLFOLKEVNAKRGVPADQLRVIFAGKELRNDMTVQNC 60
QY 61 LDQOSIVHIVORPMRKGEMNATGDDPRNNAAGCEREPOSTLRVDSSTVLPGDSVGLA 120
DB 61 LDQOSIVHIVORPMRKGEMNATGDDPRNNAAGCEREPOSTLRVDSSTVLPGDSVGLA 120
QY 121 VILHTDSRSDSPAGSPAGRSITNSFYCYCKGPCQORVOPGKLRVQSTCRQATLTLTQGP 180
DB 121 VILHTDSRSDSPAGSPAGRSITNSFYCYCKGPCQORVOPGKLRVQSTCRQATLTLTQGP 180
QY 181 SCMDVLLIPNRMSEGCSPHCPGTSABEFFPKCAHPTSDKETPVVALHLIATNSRNTTCT 240
DB 181 SCMDVLLIPNRMSEGCSPHCPGTSABEFFPKCAHPTSDKETPVVALHLIATNSRNTTCT 240
QY 241 CTDVRSPLVLFQCNRSRHVITCLDCFHLVCYTRLNDRFVHDPOLGYSILPCVAGCPNLSLKE 300
DB 241 CTDVRSPLVLFQCNRSRHVITCLDCFHLVCYTRLNDRFVHDPOLGYSILPCVAGCPNLSLKE 300
QY 301 LHHFRILGEBQYNNRYOYGABECVLOMGVLCPRPGCAGLLPEPDQRKVTCEGNGLGC 360
DB 301 LHHFRILGEBQYNNRYOYGABECVLOMGVLCPRPGCAGLLPEPDQRKVTCEGNGLGC 360
QY 361 GFACRECKEAYHGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKTTKPCPR 420
DB 361 GFACRECKEAYHGECSAVFEASGTTTQAYRVDERAAEQARWEAASKETIKTTKPCPR 420
QY 421 CHVPEKNGGCMHMKCPQPCRLIEWCNCGEEMNRVCMGDHMFV 465
DB 421 CHVPEKNGGCMHMKCPQPCRLIEWCNCGEEMNRVCMGDHMFV 465

RESULT 6
ABO07157

ID ABO07157 standard; protein; 437 AA.

XX ABO07157;

DT 13-AUG-2003 (first entry)

XX Human p53 modifying protein, SEQ ID 117.

XX Human; p53 modifier; cytostatic; cancer; cytostatic; antiangiogenic;
XX antiapoptotic; p53 pathway; breast cancer; colon cancer; kidney cancer;
XX lung cancer; ovarian cancer; angiogenesis; cell cycle;
XX apoptotic disorder; cell proliferation disorder.

XX Homo sapiens.

XX WO200239122-A1.
 PN
 XX 12-DEC-2002.
 PD
 XX
 XX 03-JUN-2002; 2002WO-US017382.
 PF
 XX 05-JUN-2001; 2001US-0296076P.
 PR 10-OCT-2001; 2001US-0328605P.
 PR 15-FEB-2002; 2002US-0357253P.
 XX
 XX (EXEL-) EXELIXIS INC.
 PA
 PI Friedland L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
 XX
 XX WPI, 2003-156859/15.
 DR N-PSDB; ACD13333.
 XX
 PT Identifying modulators of the p53 pathway for use in treating apoptotic
 PT or cell proliferation disorders, comprises screening for agents that
 PT modulate activity of a human ortholog of genes that modify the p53
 PT pathway in Drosophila.

Example 2; Page 401-402; 678bp; English.

CC The invention relates to identifying (M1) a candidate p53 pathway
 CC modulating agent, by contacting an assay system comprising a purified HM
 CC polypeptide (human orthologue or genes that modify the p53 pathway in
 CC Drosophila) or nucleic acid with a test agent under conditions, where but
 CC for the presence of the test agent, the system provides a reference
 CC activity, and detecting a test agent-biased activity of the assay system.
 CC Also included are modulating (M2) a p53 pathway of a cell, (comprising
 CC contacting a cell defective in p53 function with a candidate modulator
 CC that specifically binds to a HM polypeptide comprising an HM amino acid
 CC sequence, where p53 function is restored), modulating (M3) a p53 pathway
 CC in a mammalian cell (comprising contacting the cell with an agent that
 CC specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
 CC a disease in a patient (comprising: (a) obtaining a biological sample
 CC from the patient; (b) contacting the sample with a probe for HM
 CC expression; (c) comparing the results with a control; and (d) determining
 CC whether the comparison indicates a likelihood disease). (M1) is useful
 CC for identifying modulators of the p53 pathway. A probe for HM expression
 CC is useful for diagnosing breast, colon, kidney, lung and ovarian cancer.
 CC In a patient, where the cancer has greater than 25 % expression level.
 CC Modulators identified by (M1) are useful in a variety of diagnostic and
 CC therapeutic applications, where disease or disorder prognosis is related
 CC to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
 CC proliferation disorders (e.g. cancer). Another two new methods (M2 and
 CC M3) are useful for modulating the p53 pathway of a cell, thus restoring
 CC the p53 function of the cell, so that the cell undergoes normal
 CC proliferation or progression through the cell cycle. (M2) and (M3) are
 CC also useful for treating defects in the p53 pathway such as angiogenic,
 CC apoptotic or cell proliferation disorders. The present sequence
 CC represents a human p53 pathway modifying protein

SO Sequence 437 AA;

Query Match 92.6%; Score 2405; DB 6; Length 437;

Best Local Similarity 94.0%; Pred. No. 1.3e-202;

Matches 437; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

QY 1 MIVFPRNSHGPVVDSDTSLPOLKEVAVKQGVAPADOLVTFKGLRDMWQONCD 60
 DB 1 MIVFPRNSHGPVVDSDTSLPOLKEVAVKQGVAPADOLVTFKGLRDMWQONCD 60
 QY 61 LDQGSIVHIVQRPWKGQEMNATGDDPNNAAGCCEREPQSLTRVDI,SSSVLP,PGDSVGLA 120
 DB 61 LDQGSIVHIVQRPWKGQEMNATGDDPNNAAGCCEREPQSLTRVDI,SSSVLP,PGDSVGLA 120
 QY 121 VILHTDSRKDSSPAGSPAGRSIYNSFYVCKGPGQVQVQPKLRVQGSTRCQATLITGQP 180
 DB 121 VILHTDSRKDSSPAGSPAGRSIYNSFYVCKGPGQVQVQPKLRVQGSTRCQATLITGQP 180

QY 181 SCMDVLIJPNRNSGBCQSPHCPTSAEFPFKCGAHPDSDKETPVALLIATNSRNTICT 240
 DB 179 -----EFPFKCGAHPDSDKETPVALLIATNSRNTICT 212
 QY 241 CTDVRSPLYVFOGNSRHYICLDCFHLXCTTRLANDRQFVNDPQLGYSLPVACGPNLIE 300
 DB 213 CTDVRSPLYVFOGNSRHYICLDCFHLXCTTRLANDRQFVNDPQLGYSLPVACGPNLIE 272
 QY 301 LHHFRILGEOYNNRYQOYGAEECVLQMGVLCPRPGCGAGLLPEPDQKVTCEGNGLGC 360
 DB 273 LHHFRILGEOYNNRYQOYGAEECVLQMGVLCPRPGCGAGLLPEPDQKVTCEGNGLGC 332
 QY 361 GFAPCRCKEAYHEGSCSAVFASGTTTQAYRVDERRAQAQWEAASKETIKTTPCR 420
 DB 333 GFAPCRCKEAYHEGSCSAVFASGTTTQAYRVDERRAQAQWEAASKETIKTTPCR 392
 QY 421 CHVPVEKXGSCMHMKCPPOQCRLENCWNGCGEMNRYCMGDHFDV 465
 DB 393 CHVPVEKXGSCMHMKCPPOQCRLENCWNGCGEMNRYCMGDHFDV 437

RESULT 7

AAB67517
 ID AAB67517 standard; protein; 464 AA.

XX AAB67517;
 AC

DT 29-MAY-2001 (first entry)

DE Amino acid sequence of a murine parkin2 polypeptide.

XX Parkinson's disease; 6q25.2-27; neurodegenerative disease;

KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;

KW Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;

KW brain tumour; head trauma; stroke; vascular irregularity;

XX metabolic irregularity.

OS Mus sp.

XX EP1081225-A1.

XX 07-MAR-2001.

XX 30-AUG-1999; 99EP-00116766.

XX 30-AUG-1999; 99EP-00116766.

XX (BIOF-) BIOFRONTIERA PHARM GMBH.

XX Luebbert H;

XX WPI; 2001-212797/22.

XX N-PSDB; AAF55244.

XX New polynucleotides encoding mouse parkin2 protein, useful for producing

PT a transgenic non-human animal as an animal model for neurodegenerative

PT diseases.

XX Disclosure; Page 17-19; 62pp; English.

XX The present sequence represents a murine parkin2 polypeptide. Mutations

CC or deletions in the parkin2 gene cause Parkinson's disease in humans. The

CC human parkin2 gene is located in gene region 6q25.2-27. Parkinson

CC polypeptides and polynucleotides are useful for testing the efficacy

CC neurodegenerative diseases. They are also useful for testing the efficacy

CC of the treatment of a neurodegenerative disease such as Parkinson's

CC disease, Alzheimer's disease, Huntington's disease, amyotrophic lateral

CC sclerosis, Multi-system atrophy, Wilson's disease, Pick's disease, Prion

CC disease, and secondary causes inducing Parkinson's syndromes like toxins,

CC drugs, brain tumours, head trauma, stroke, vascular irregularities or

CC metabolic irregularities, associated with a less active or non-active

CC parkin protein

Sequence 464 AA;

| Query Match | Score | DB 4 | Length |
|-----------------------|---------|---------|--------|
| 83.6%; | 2170.5; | 464; | |
| Best Local Similarity | 83.4%; | 2169.5; | 464; |

Matches 388; Conservative 29; Mismatches 47; Indels 1.

1 MY/EVBENCUCEDUMMOMATNOT.....

| | | | | |
|----|-----|----------------------------|---|-----|
| QY | 1 | MIVFRRFSSHGFEVENDSTSI | FOLEKAVARQSGVPAOLAVITRAGKELRNDMTVQNC | 60 |
| Db | 1 | MIVVRNRSSGFEVENDSTSI | LLOLEKAVARQSGVPAOLAVITRAGKELRNDMTVQNC | 60 |
| QY | 61 | LDQOSIVHIVORPWRKQEMNATGAD | PRNAAGCEREPOSILTRVDSSSVLPBGDSVLA | 120 |
| Db | 61 | LEOOSIVHIVORPWRKSHETNAGSG | EDPOSTSEGSIMWERSILTRVDSLSHLLPVDVGLA | 120 |
| QY | 121 | VIIHTDSRKSPRAGSPAGRSINSTVY | CKBPQORVOPGKIRVOCSTCROATITLTQSP | 180 |
| Db | 121 | VIIHTDSRKSEARGVV-KPTNSFFI | CKGCHVOPGKIRVOCSTCROATITLTQSP | 179 |
| QY | 181 | SCMDVLIIPRMSGECOSPHCPGTS | AEFFKGAHPTSDKXENPVAMHIANNSNITCIT | 240 |
| Db | 180 | SCMDVLIIPRMSGECOSPHCPETRA | EFFKGAHPTSDKSTVAMLNITNSRISPCIA | 239 |
| QY | 241 | CTDVRSGVIVFOCNSRHVITCLD | FHLVCYTRILNDRQVPHNDOLGYSILPCVAGCNSLIKE | 300 |
| Db | 240 | CTDVRSGVIVFOCNRHHRVITCLD | FHLVCYTRILNDRQVPHNDOLGYSILPCVAGCNSLIKE | 299 |
| QY | 301 | LHHFRILGEEQYNYRVOYGAEBGV | LOMGVILCPRRGCCAGILPBPDOAKYTCBSGNGLG | 360 |
| Db | 300 | LHHFRILGEEQYTRYQYGAEBGV | LOMGVILCPRRGCCAGILPBPDOGRKYTCBSGNGLG | 359 |
| QY | 361 | GFAPFCEKEAVYHEGCSAFVFA | SGTITTOYAVVDERRAALQARMBAASKETIYKKTTRCPR | 420 |
| Db | 360 | GFVFCDCDEAVYHEGDCSLLEP | SGATSGOAVRVDBRAALQARMBAASKETIYKKTTRCPR | 419 |
| QY | 421 | CHVPEKNGCMHMKPOPCRLFL | ENCMNGCGEMNRCVCMGDHMDV | 465 |
| Db | 420 | CNVPLEKNGCMHMKPOPCRLFL | ENCMNGCGEMNRCVCMGDHMDV | 464 |

RESULT 8

| ID | standard; protein; 464 AA. |
|----------|----------------------------|
| AA067531 | |
| AA067532 | |
| AA067533 | |
| AA067534 | |
| AA067535 | |
| AA067536 | |
| AA067537 | |
| AA067538 | |
| AA067539 | |
| AA067540 | |
| AA067541 | |
| AA067542 | |
| AA067543 | |
| AA067544 | |
| AA067545 | |
| AA067546 | |
| AA067547 | |
| AA067548 | |
| AA067549 | |
| AA067550 | |
| AA067551 | |
| AA067552 | |
| AA067553 | |
| AA067554 | |
| AA067555 | |
| AA067556 | |
| AA067557 | |
| AA067558 | |
| AA067559 | |
| AA067560 | |
| AA067561 | |
| AA067562 | |
| AA067563 | |
| AA067564 | |
| AA067565 | |
| AA067566 | |
| AA067567 | |
| AA067568 | |
| AA067569 | |
| AA067570 | |
| AA067571 | |
| AA067572 | |
| AA067573 | |
| AA067574 | |
| AA067575 | |
| AA067576 | |
| AA067577 | |
| AA067578 | |
| AA067579 | |
| AA067580 | |
| AA067581 | |
| AA067582 | |
| AA067583 | |
| AA067584 | |
| AA067585 | |
| AA067586 | |
| AA067587 | |
| AA067588 | |
| AA067589 | |
| AA067590 | |
| AA067591 | |
| AA067592 | |
| AA067593 | |
| AA067594 | |
| AA067595 | |
| AA067596 | |
| AA067597 | |
| AA067598 | |
| AA067599 | |
| AA067600 | |
| AA067601 | |
| AA067602 | |
| AA067603 | |
| AA067604 | |
| AA067605 | |
| AA067606 | |
| AA067607 | |
| AA067608 | |
| AA067609 | |
| AA067610 | |
| AA067611 | |
| AA067612 | |
| AA067613 | |
| AA067614 | |
| AA067615 | |
| AA067616 | |
| AA067617 | |
| AA067618 | |
| AA067619 | |
| AA067620 | |
| AA067621 | |
| AA067622 | |
| AA067623 | |
| AA067624 | |
| AA067625 | |
| AA067626 | |
| AA067627 | |
| AA067628 | |
| AA067629 | |
| AA067630 | |
| AA067631 | |
| AA067632 | |
| AA067633 | |
| AA067634 | |
| AA067635 | |
| AA067636 | |
| AA067637 | |
| AA067638 | |
| AA067639 | |
| AA067640 | |
| AA067641 | |
| AA067642 | |
| AA067643 | |
| AA067644 | |
| AA067645 | |
| AA067646 | |
| AA067647 | |
| AA067648 | |
| AA067649 | |
| AA067650 | |
| AA067651 | |
| AA067652 | |
| AA067653 | |
| AA067654 | |
| AA067655 | |
| AA067656 | |
| AA067657 | |
| AA067658 | |
| AA067659 | |
| AA067660 | |
| AA067661 | |
| AA067662 | |
| AA067663 | |
| AA067664 | |
| AA067665 | |
| AA067666 | |
| AA067667 | |
| AA067668 | |
| AA067669 | |
| AA067670 | |
| AA067671 | |
| AA067672 | |
| AA067673 | |
| AA067674 | |
| AA067675 | |
| AA067676 | |
| AA067677 | |
| AA067678 | |
| AA067679 | |
| AA067680 | |
| AA067681 | |
| AA067682 | |
| AA067683 | |
| AA067684 | |
| AA067685 | |
| AA067686 | |
| AA067687 | |
| AA067688 | |
| AA067689 | |
| AA067690 | |
| AA067691 | |
| AA067692 | |
| | |

AC AAB67531;

DT 29-MAY-2001 (first entry)

Amino acid sequence of a mutated murine parkin2 polypeptide.

NM Parkinson's disease, 6q25.2-27; neurodegenerative disease;
 NM Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
 NM Multi-system atrophy; Wilson's disease; Pick's disease;
 NM brain tumour; head trauma; stroke; vascular irregularity;
 NM metabolic irregularity.

OS Mus & K

PN EP1081225-A1.

PD 07-MAR-2001.

30-AUG-1999; 99EP-00116766.

PR 30-AUG-1999; 99EP-00116766.
XX

PA (BIOF-) BIOFRONTERA PHARM GMBH.
XY

Luebbert H;

DR WPI; 2001-212797/22.
DR N-PSDB; AAF55258.

DR N-PSDB; AAF55258.

1 New polynucleotides encoding mouse parkin2 protein, useful for producing

PT a transgenic non-human animal as an animal model for neurodegenerative PT diseases.

PT diseases.

PS Claim 7; Page 47-49; 62pp; English.

The present sequence represents a murine parkin polypeptide. The sequence contains the mutation Lys161Asn. Mutations or deletions in the parkin2 gene cause Parkinson's disease in humans. The human parkin2 gene is located in gene region 6q25.2-27. Parkin2 polypeptides and polynucleotides are useful for analysing neurodegenerative diseases. They are also useful for testing the efficacy of the treatment of a neurodegenerative disease such as Parkinson's disease, Alzheimer's disease, Huntington's disease, amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease, Pick's disease, Prion disease, and secondary causes including Parkinson's syndromes like toxins, drugs, brain tumours, head trauma, stroke, vascular irregularities or metabolic irregularities, associated with a less active or non-active parkin protein

SQ Sequence 464 AA;

Query Match 83.48; Score 2165.5; DB 4; Length 464;

Best Local Similarity 83.2%; Pred. No. 1.7e-181;
Matches 287; Consistency 30. We used

Matches 36; conservative 29; mismatches 48; indels 1; gaps 1;

| | | | | | | | | | |
|----|-----|-----------------------|---------------|------------|------------|------------|-----------|------------------|--------------|
| QY | 1 | MIVFRRFSSHGPFVEVDSTSI | FOLEEVAKRGQ | PAPOLRIT | FAFGKEL | RNDMT | VQND | 60 | |
| Db | 1 | MIVFRRFSSYGFVEVDSTSI | LOLKEVAKRGQ | SPADQLKVI | PAFGKEL | PHMLT | VQND | 60 | |
| QY | 61 | LDQSIYHIVOR | PMRKQEMNATGSD | PRNNAAGC | GEREP | OSTLR | VDLSSVLP | PGDVGLA 120 | |
| Db | 61 | LEQSIYHIVOR | PMRRRHETN | ASGDEPOST | SGSIWESRSL | ITLR | VDLSHHTLP | PVDSVGLA 120 | |
| QY | 121 | VIIHATDRKDS | PPAGSPAGRSI | YNSPVYVCKE | QCRVQBGXLR | YOCSTCRQA | TLTLTGCP | 180 | |
| Db | 121 | VIIHATDRKDS | PPAGSARGV | KPTYNSEFII | CKGPKH | QVQGNLR | AVOCGTCKO | ATLTLAAGP 179 | |
| QY | 181 | SCMDVLI | IPNRMSGECOS | PHCPGTS | AEFFPKCGA | FTSDKET | PVALMHI | IATNSRNITCIT 240 | |
| Db | 180 | SCMDVLI | IPNRMSGECOS | PDPCGT | RAEPF | FGCAHPT | SDKDT | SVALNLTITSNR | RI PCIA 239 |
| QY | 241 | CTDVRSPVL | AFQCNRRHVI | CLDCEHLY | YCTTRLND | QPHADPOL | GLSYLSP | CVAGC | PNLSLIXE 300 |
| Db | 240 | CTDVRSPVL | AFQCNHRRHVI | CLDCEHLY | YCTTRLND | QPHADPOL | GLSYLSP | CVAGC | PNLSLIXE 299 |
| QY | 301 | LIHFRILIGE | ROYNRYOQY | GAEBECVLO | MGVGLCP | RRGCGAGALL | PEPDOR | KMTCEG | NGILGC 360 |
| Db | 300 | LIHFRILIGE | ROYNRYOQY | GAEBECVLO | MGVGLCP | RRGCGAGALL | PEPDQAKVT | CEG | NGILGC 359 |
| QY | 361 | GFAFCBCE | KEAHHEGCS | AVFEASGTTT | QAVR | VDERRAAE | QARWEA | SKETIK | TKTKPCPR 420 |
| Db | 360 | GFAFCBCE | KEAHHEGCS | ILTEPS | GATQAY | VRDRALE | QARWEA | SKETIK | TKTKCPR 419 |
| QY | 421 | CHAFVE | NGGCMHMK | CDPOQCR | LEMG | CNNGCE | EMNR | VCMG | GDWPFV 465 |
| Db | 420 | CNVPI | ERNGGCMHMK | CDPOQCR | LEMG | CNNGCE | EMNNA | CGMD | WPFV 464 |

RESULT 9

AAB67532 standard; protein; 464 AA

AAC AAB67532;

29-MAY-2001 (first entry)

Amino acid sequence of a mutated murine parkin2 polypeptide.

M Alzheimer's disease; 6625 2-27; neurodegenerative disease.
 M Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis.
 M Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease.
 M brain tumour; head trauma; stroke; vascular irregularity; metabolic irregularity.

XX Mus sp.
XX EPI081225-A1.
XX 07-MAR-2001.
XX 30-AUG-1999; 99EP-00116766.
XX 30-AUG-1999; 99EP-00116766.
XX (BIOF-) BIOFRONTIERA PHARM GMBH.
XX Luebbert H;
XX WPI: 2001-212797/22.
XX N-PSDB; AAF55259.
XX New polynucleotides encoding mouse parkin2 protein, useful for producing a transgenic non-human animal as an animal model for neurodegenerative diseases.
XX Claim 7, Page 49-51; 62pp; English.
XX The present sequence represents a murine parkin2 polypeptide. The sequence contains the mutation Thr415Asn. Mutations or deletions in the parkin2 gene cause Parkinson's disease in humans. The human parkin2 gene is located in gene region 6q25.2-27. Parkin2 polypeptides and polynucleotides are useful for analysing neurodegenerative diseases. They are also useful for testing the efficacy of the treatment of a neurodegenerative disease such as Parkinson's disease, Alzheimer's disease, Huntington's disease, amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease, Pick's disease, Prion disease, and secondary causes inducing Parkinson's syndromes like toxins, drugs, brain tumours, head trauma, stroke, vascular irregularities or metabolic irregularities, associated with a less active or non-active parkin protein
XX Sequence 464 AA;
SQ
Query Match 83.4%; Score 2165.5; DB 4; Length 464;
Best Local Similarity 83.2%; Pred. No. 1.7e-181;
Matches 387; Conservative 29; Mismatches 48; Indels 1; Gaps 1;
QY 1 MIVFVRNSHGFPVEVDSDTISIFOLKEVAVKRGVPAQQLVIFAGKELRNDWTQNC 60
DB 1 MIVFVRNSHGFPVEVDSDTISIFOLKEVAVKRGVPAQQLVIFAGKELRNDWTQNC 60
QY 61 LDQOSIVHIVQPRRSHETNAGSGDEPOSTSGSIWESRSLTRVDSLHTLPVDSVGLA 120
DB 61 LDQOSIVHIVQPRRSHETNAGSGDEPOSTSGSIWESRSLTRVDSLHTLPVDSVGLA 120
QY 121 VILHTDSRKDPPASPGRSIYNSFYVCKGPCORVQBGKLRVQCSSTGROATLTTCGP 180
DB 121 VILHTDSRKDPPASPGRSIYNSFYVCKGPCORVQBGKLRVQCSSTGROATLTTCGP 180
QY 121 VILHTDSRKDPPASPGRSIYNSFYVCKGPCORVQBGKLRVQCSSTGROATLTTCGP 180
DB 121 VILHTDSRKDPPASPGRSIYNSFYVCKGPCORVQBGKLRVQCSSTGROATLTTCGP 180
QY 181 SCWDVVLIPNRMSGEGQSPHCPGTSNAFFKCGAHPTSKETPPVLAHLATNSRNTCTT 240
DB 181 SCWDVVLIPNRMSGEGQSPHCPGTSNAFFKCGAHPTSKETPPVLAHLATNSRNTCTT 240
QY 241 CTDVRSPTLVPCQNSRHVCLDPCFHLVCTRLNDROFVNDPOLGYSILPCVACGPNLSLIE 300
DB 241 CTDVRSPTLVPCQNSRHVCLDPCFHLVCTRLNDROFVNDPOLGYSILPCVACGPNLSLIE 300
QY 301 LHHFRILGEGYNNRYOQYGAEBGVLMGGVLCPRPGCAGLLPEPDQKVTCEGGNGLGC 360
DB 301 LHHFRILGEGYNNRYOQYGAEBGVLMGGVLCPRPGCAGLLPEPDQKVTCEGGNGLGC 360
QY 361 GFAFCRECKEAYHEGECASVFEASGTTTQAYRVDEAAAOARWEAASKTITKTTKPCR 420
DB 361 GFAFCRECKEAYHEGECASVFEASGTTTQAYRVDEAAAOARWEAASKTITKTTKPCR 420
QY 421 CHVPVEKINGCMHMKCPQPCCKLEWCMNCGCEWNRVCMGDHWFVDV 465
DB 421 CHVPVEKINGCMHMKCPQPCCKLEWCMNCGCEWNRVCMGDHWFVDV 465

DB 420 CHVPVEKINGCMHMKCPQPCCKLEWCMNCGCEWNRVCMGDHWFVDV 464
RESULT 10
ID AAB67533
ID AAB67533 standard; protein; 451 AA.
AC AAB67533;
DT 29-MAY-2001 (first entry)
XX Amino acid sequence of a mutated murine parkin2 polypeptide.
XX Parkinson's disease; 6q25.2-27; neurodegenerative disease;
XX Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
XX Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
XX brain tumour; head trauma; stroke; vascular irregularity;
XX metabolic irregularity.
OS Mus sp.
XX EPI081225-A1.
XX 07-MAR-2001.
XX 30-AUG-1999; 99EP-00116766.
XX 30-AUG-1999; 99EP-00116766.
XX (BIOF-) BIOFRONTIERA PHARM GMBH.
XX Luebbert H;
XX WPI: 2001-212797/22.
XX N-PSDB; AAF55260.
XX New polynucleotides encoding mouse parkin2 protein, useful for producing a transgenic non-human animal as an animal model for neurodegenerative diseases.
XX Claim 7, Page 51-53; 62pp; English.
XX The present sequence represents a murine parkin2 polypeptide. The sequence contains the mutation Tyr453Stop. Mutations or deletions in the parkin2 gene cause Parkinson's disease in humans. The human parkin2 gene is located in gene region 6q25.2-27. Parkin2 polypeptides and polynucleotides are useful for analysing neurodegenerative diseases. They are also useful for testing the efficacy of the treatment of a neurodegenerative disease such as Parkinson's disease, Alzheimer's disease, Huntington's disease, amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease, Pick's disease, Prion disease, and secondary causes inducing Parkinson's syndromes like toxins, drugs, brain tumours, head trauma, stroke, vascular irregularities or metabolic irregularities, associated with a less active or non-active parkin protein
XX Sequence 451 AA;
SQ
Query Match 80.4%; Score 2087.5; DB 4; Length 451;
Best Local Similarity 83.2%; Pred. No. 1.2e-174;
Matches 376; Conservative 29; Mismatches 46; Indels 1; Gaps 1;
QY 1 MIVFVRNSHGFPVEVDSDTISIFOLKEVAVKRGVPAQQLVIFAGKELRNDWTQNC 60
DB 1 MIVFVRNSHGFPVEVDSDTISIFOLKEVAVKRGVPAQQLVIFAGKELRNDWTQNC 60
QY 61 LDQOSIVHIVQPRRSHETNAGSGDEPOSTSGSIWESRSLTRVDSLHTLPVDSVGLA 120
DB 61 LDQOSIVHIVQPRRSHETNAGSGDEPOSTSGSIWESRSLTRVDSLHTLPVDSVGLA 120
QY 121 VILHTDSRKDPPASPGRSIYNSFYVCKGPCORVQBGKLRVQCSSTGROATLTTCGP 180
DB 121 VILHTDSRKDPPASPGRSIYNSFYVCKGPCORVQBGKLRVQCSSTGROATLTTCGP 180

Db 121 VILDTDSKRDSEANAGPV-KFTYNSFFIYCKGPKCHKVOPGKLRVQCCTCKOATVTLAAGP 179
 Qy 181 SCMDVLLIPNRMSEGCSPHCPGTSAPFPKGAHPTSDKETPVVALHLIATNSRNTTCT 240
 Db 180 SCMDVLLIPNRMSEGCSPHCPGTSAPFPKGAHPTSDKETPVVALHLIATNSRNTTCT 239
 Qy 241 CTDVASPVLVFQCSNRHVICLDCEHLVCTRLNDROFVHDPOLGYSLPVAGCPNSLIKE 300
 Db 240 CTDVASPVLVFQCSNRHVICLDCEHLVCTRLNDROFVHDPOLGYSLPVAGCPNSLIKE 299
 Qy 301 LHHFRILIGBEOYNRVQOYGAEECTLOMGVLCPRPGCAGLLPDPDQKVTCEGNGLGC 360
 Db 300 LHHFRILIGBEOYNRVQOYGAEECTLOMGVLCPRPGCAGLLPDPDQKVTCEGNGLGC 359
 Qy 361 GFAPCRCKEAYHGECSAVFEASGTTTQAYRVDERAARWEAKSKETIKTTTKECPR 420
 Db 360 GFAPCRCKEAYHGECSAVFEASGTTTQAYRVDERAARWEAKSKETIKTTTKECPR 419
 Qy 421 CHVEVERKNGGCMHMKCPQPCRLMCMNCCGE 452
 Db 420 CNVEIERKNGGCMHMKCPQPCRLMCMNCCGE 451

RESULT 11

ABO07158
 ID ABO07158 standard; protein; 316 AA.

AC ABO07158;
 DT 13-AUG-2003 (first entry)

DE Human p53 modifying protein, SEQ ID 118.

KM Human; p53 modifier; cytosolic; cancer; cytosolic; antiangiogenic;
 KM antiapoptotic; p53 pathway; breast cancer; colon cancer; kidney cancer;
 KM lung cancer; ovarian cancer; angiogenesis; cell cycle;
 KM apoptotic disorder; cell proliferation disorder.

OS Homo sapiens.

PN W0200299122-A1.

PD 12-DEC-2002.

PE 03-JUN-2002; 2002MO-US017382.

PR 05-JUN-2001; 2001US-0296076P.

PR 10-OCT-2001; 2001US-0328605P.

PR 15-FEB-2002; 2002US-0357253P.

PA (EXEL-) EXELIXIS INC.

PI Friedman L, Plowman GD, Belvin M, Francis-Liang H, Li D, Funke RP;

DR WPI; 2003-156859/15.

DR N-PSDB; ACD13334.

PT Identifying modulators of the p53 pathway for use in treating apoptotic
 PT modulate activity of a human ortholog of genes that modify the p53
 PT pathway in Drosophila.

PS Example 2; Page 402-403; 678bp; English.

CC The invention relates to identifying (M1) a candidate p53 pathway
 CC modulating agent, by contacting an assay system comprising a purified HM
 CC polypeptide (human ortholog of genes that modify the p53 pathway in
 CC Drosophila) or nucleic acid with a test agent under conditions, where but
 CC for the presence of the test agent, the system provides a reference
 CC activity, and detecting a test agent-biased activity of the assay system.
 CC Also included are modulating (M2) a p53 pathway of a cell (comprising
 CC contacting a cell defective in p53 function with a candidate modulator
 CC that specifically binds to a HM polypeptide comprising an HM amino acid

CC sequence, where p53 function is restored), modulating (M3) a p53 pathway
 CC in a mammalian cell (comprising contacting the cell with an agent that
 CC specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
 CC a disease in a patient (comprising: (a) obtaining a biological sample
 CC from the patient; (b) contacting the sample with a probe for HM
 CC expression; (c) comparing the results with a control; and (d) determining
 CC whether the comparison indicates a likelihood disease). (M1) is useful
 CC for identifying modulators of the p53 pathway. A probe for HM expression
 CC is useful for diagnosing breast, colon, kidney, lung and ovarian cancer.
 CC In a patient, where the cancer has greater than 25 % expression level.
 CC Modulators identified by (M1) are useful in a variety of diagnostic and
 CC therapeutic applications, where disease or disorder prognosis is related
 CC to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
 CC proliferation disorders (e.g. cancer). Another two new methods (M2 and
 CC M3) are useful for modulating the p53 pathway of a cell, thus restoring
 CC the p53 function of the cell, so that the cell undergoes normal
 CC proliferation or progression through the cell cycle. (M2) and (M3) are
 CC also useful for treating defects in the p53 pathway such as angiogenic,
 CC apoptotic or cell proliferation disorders. The present sequence
 CC represents a human p53 pathway modifying protein

SO Sequence 316 AA;

Query Match 65.4%; Score 1696.5; DB 6; Length 316;
 Best Local Similarity 68.0%; Pred. No. 1.8e-140;
 Matches 316; Conservative 0; Mismatches 0; Indels 149; Gaps 1;

Qy 1 MIVFVRENSSHGFPVNDSDTSIFOLKEVNAKRGVPAADQLRVTFAGKELRNDVTQVNC 60
 Db 1 MIVFVRENSSHGFPVNDSDTSIFOLKEVNAKRGVPAADQLRVTFAGKELRNDVTQV--- 57
 Qy 61 LDQGSIVHIVQPRKGEENATGDDPRNAAGCEREPSLTVNDLSSVLPEDSVGLA 120
 Db 58 ----- 57
 Qy 121 VILHTRSKDSPAGSPAGRSYNSFYVYCKGQVQVOPGKLRVQCSTCRQATLTLTQGP 180
 Db 58 ----- 57
 Qy 181 SCMDVLLIPNRMSEGCSPHCPGTSAPFPKGAHPTSDKETPVVALHLIATNSRNTTCT 240
 Db 58 -----EFPKGAHPTSDKETPVVALHLIATNSRNTTCT 91
 Qy 241 CTDVASPVLVFQCSNRHVICLDCEHLVCTRLNDROFVHDPOLGYSLPVAGCPNSLIKE 300
 Db 92 CTDVASPVLVFQCSNRHVICLDCEHLVCTRLNDROFVHDPOLGYSLPVAGCPNSLIKE 151
 Qy 301 LHHFRILIGBEOYNRVQOYGAEECTLOMGVLCPRPGCAGLLPDPDQKVTCEGNGLGC 360
 Db 152 LHHFRILIGBEOYNRVQOYGAEECTLOMGVLCPRPGCAGLLPDPDQKVTCEGNGLGC 211
 Qy 361 GFAPCRCKEAYHGECSAVFEASGTTTQAYRVDERAARWEAKSKETIKTTTKECPR 420
 Db 212 GFAPCRCKEAYHGECSAVFEASGTTTQAYRVDERAARWEAKSKETIKTTTKECPR 271
 Qy 421 CHVEVERKNGGCMHMKCPQPCRLMCMNCCGEENRVCMGDHMFV 465
 Db 272 CHVEVERKNGGCMHMKCPQPCRLMCMNCCGEENRVCMGDHMFV 316

RESULT 12

ABAB67521
 ID AAB67521 standard; protein; 344 AA.

AC AAB67521;

DT 29-MAY-2001 (first entry)

DE Amino acid sequence of a murine truncated parkin2 polypeptide.

KM Parkinson's disease; 6q25.2-27; neurodegenerative disease;
 KM Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
 KM multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;

```
KW brain tumour; head trauma; stroke; vascular irregularity;
KW metabolic irregularity.
OS Mus sp.
FN EPI081225-A1.
PD 07-MAR-2001.
XX
XX
XX 30-AUG-1999; 99EP-00116766.
XX
XX 30-AUG-1999; 99EP-00116766.
XX
XX 30-AUG-1999; 99EP-00116766.
XX
XX (BIOF-) BIOFRONTIERA PHARM GMBH.
XX
XX Luebert H;
XX
XX WPI; 2001-212797/22.
XX
XX N-PSDB; AAF55248.
XX
XX
XX New polynucleotides encoding mouse parkin2 protein, useful for producing
XX a transgenic non-human animal as an animal model for neurodegenerative
XX diseases.
XX
XX Claim 7; Page 40-41; 62pp; English.
XX
XX
XX The present sequence represents a murine parkin2 polypeptide. The
XX polynucleotide sequence contains a deletion, leading to a truncated
XX protein. Mutations or deletions in the parkin2 gene cause Parkinson's
XX disease in humans. The human parkin2 gene is located in gene region
XX 6q25.2-27. Parkinson polypeptides and polynucleotides are useful for
XX analysing neurodegenerative diseases. They are also useful for testing
XX the efficacy of the treatment of a neurodegenerative disease such as
XX Parkinson's disease, Alzheimer's disease, Huntington's disease,
XX amyotrophic lateral sclerosis, multi-system atrophy, Wilson's disease,
XX Pick's disease, Prion disease, and secondary causes inducing Parkinson's
XX syndromes like toxins, drugs, brain tumours, head trauma, stroke,
XX CC vascular irregularities or metabolic irregularities, associated with a
XX less active or non-active parkin protein
XX
XX Sequence 344 AA;
SQ
Query Match 64.7%; Score 1679.5; DB 4; Length 344;
Best Local Similarity 66.0%; Pred. No. 6,3e-139;
Matches 307; Conservative 14; Mismatches 23; Indels 121; Gaps 1;
QY 1 MIVFVRNNSHGFPVEVSDTSIFQLKEVAVARQGVADQLVIFAGKELRNDWYQNC 60
DB 1 MIVFVRNNSYGFPEVSDTSILQLKEVAVARQGVADQLVIFAGKELPHNLTV---- 56
QY 61 LDOQSIWVIVQRPWRKGGEMNATGGDDPRNNAAGGCRERPOSILTRVDLSSTVLRGDSVGLA 120
DB 57 ----- 56
QY 121 VILHTDSRKDSPAGSPAGRSIYNSFYVCKGPCQGVQPKLRVQSTCRQATLTLTGSP 180
DB 57 ----- 56
QY 57 ----- 59
DB 181 SCWDDVLIPNRMSGECQSPHCPGTSAPFFKCGAHPDSKETPVALHLIATNSRNTTCT 240
DB 60 SCWDDVLIPNRMSGECQSPDCPTRAEPFKCGAHPDTSKDTSVALLTNSRSPICIA 119
QY 241 CTDVASSPVTVFQCNSTRVYCLDCCFHLVYCTRLANDRQFVNDPQLGYSLPCVACGPNLSLIE 300
DB 120 CTDVASSPVTVFQCNSTRVYCLDCCFHLVYCTRLANDRQFVNDPQLGYSLPCVACGPNLSLIE 179
QY 301 LHHFRILGEEQYRNYQOYGAABECVLOMGVLCPRPGCAGALLPEPDCRYKTCGGAGLGC 360
DB 180 LHHFRILGEEQYRNYQOYGAABECVLOMGVLCPRPGCAGALLPEPDCRYKTCGGAGLGC 239
QY 361 GFAPCRBCKEAYHEGCSAVFEASGTTTQAYRVERDAQAQARWEAASKETIKTKTPCR 420
DB 240 GFVFCRDCKEAYHEGCDLSLLEPSGATSAAYRVDRKRAAQAARWEAASKETIKTKTPCR 299
```

```
QY 421 CHVPEKNGGCMHMKCPQPCRLIEWCMNGCEMNRYCMGDHWFV 465
DB 300 CNVPIEKNGGCMHMKCPQPCRLIEWCMNGCEMNRYCMGDHWFV 344
RESULT 13
AAB67526
ID AAB67526 standard; protein; 296 AA.
XX
XX AAB67526;
AC
XX 29-MAY-2001 (first entry)
DT
XX
XX Amino acid sequence of a murine truncated parkin2 polypeptide.
DE
XX
XX Parkinson's disease; 6q25.2-27; neurodegenerative disease;
KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
KW Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
KW brain tumour; head trauma; stroke; vascular irregularity;
KW metabolic irregularity.
XX
XX Mus sp.
XX
XX EPI081225-A1.
XX
XX PD 07-MAR-2001.
XX
XX 30-AUG-1999; 99EP-00116766.
XX
XX 30-AUG-1999; 99EP-00116766.
XX
XX (BIOF-) BIOFRONTIERA PHARM GMBH.
XX
XX Luebert H;
XX
XX WPI; 2001-212797/22.
XX
XX N-PSDB; AAF55253.
XX
XX New polynucleotides encoding mouse parkin2 protein, useful for producing
XX a transgenic non-human animal as an animal model for neurodegenerative
XX diseases.
XX
XX Claim 7; Page 44-45; 62pp; English.
XX
XX
XX The present sequence represents a murine parkin2 polypeptide. The
XX polynucleotide sequence contains a deletion, leading to a truncated
XX protein. Mutations or deletions in the parkin2 gene cause Parkinson's
XX disease in humans. The human parkin2 gene is located in gene region
XX 6q25.2-27. Parkinson polypeptides and polynucleotides are useful for
XX analysing neurodegenerative diseases. They are also useful for testing
XX the efficacy of the treatment of a neurodegenerative disease such as
XX Parkinson's disease, Alzheimer's disease, Huntington's disease,
XX amyotrophic lateral sclerosis, multi-system atrophy, Wilson's disease,
XX Pick's disease, Prion disease, and secondary causes inducing Parkinson's
XX syndromes like toxins, drugs, brain tumours, head trauma, stroke,
XX CC vascular irregularities or metabolic irregularities, associated with a
XX less active or non-active parkin protein
XX
XX Sequence 296 AA;
SQ
Query Match 47.6%; Score 1234.5; DB 4; Length 296;
Best Local Similarity 80.0%; Pred. No. 7,3e-100;
Matches 232; Conservative 20; Mismatches 37; Indels 1; Gaps 1;
QY 1 MIVFVRNNSHGFPVEVSDTSIFQLKEVAVARQGVADQLVIFAGKELRNDWYQNC 60
DB 1 MIVFVRNNSYGFPEVSDTSILQLKEVAVARQGVADQLVIFAGKELPHNLTVQNC 60
QY 61 LDOQSIWVIVQRPWRKGGEMNATGGDDPRNNAAGGCRERPOSILTRVDLSSTVLRGDSVGLA 120
DB 61 LDOQSIWVIVQRRRSHETNAGSDDEPOSTEGSIEWERSLTRVDLSHTLPVSDVGLA 120
```

QY 121 VIHTSRKSPSPAGSPAGSINSFYVYCKGQCRVQSGKRLRVQSGTCRQATLTTLTQGP 180
DB 121 VIHTSRKSPSPAGSPAGSINSFYVYCKGQCRVQSGKRLRVQSGTCRQATLTTLTQGP 179
QY 181 SCMDVLIIPRMGECQSPHCPGTSAPFFPKCGAHPRTSDKETPVALHLIATNSRNTCTT 240
DB 180 SCMDVLIIPRMGECQSPHCPGTSAPFFPKCGAHPRTSDKETPVALHLIATNSRNTCTT 239
QY 241 CTDVRSPLVLFQCNRRHVICLDCEHLYCVTRLNDQFVHDPOLGYSILPCV 290
DB 240 CTDVRSPLVLFQCNRRHVICLDCEHLYCVTRLNDQFVHDPOLGYSILPCV 289

RESULT 14

ABB65114
ID ABB65114 standard; protein; 468 AA.

AC ABB65114;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 22134.

KM Drosophila; developmental biology; cell signalling; insecticide, pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PMD, Myers EW;

DR WPI; 2001-656860/75.

XX N-PSTDB; ABL09217.

PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.

PS Disclosure; SEQ ID NO 22134; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceuticals. The invention discloses genomic DNA sequences (AB16176-AB130511), expressed DNA sequences (AB101840-AB16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 468 AA;

Query Match 40.8%; Score 1058.5; DB 4; Length 468;

Best Local Similarity 42.6%; Pred. No. 4e-84;

Matches 201; Conservative 80; Mismatches 158; Indels 33; Gaps 9;

QY 3 VVVRNSHGPFVAVDSTSIPLQLEKVAKRGVADQLRVTFAGKELRNDWTVMQCDLD 62

DB 18 IYKNTNTGKTLTVNLEPQWDIKNVKELVAPQLGLQDDLKIIFAKELSDATTEBQDLG 77

QY 63 QOSIVHIVQ-RPMRKGQEMNATGDDPNNAAAGCGRERQSLTRVULSSVLPGBSVGLAV 121

DB 78 QOSVLIHAIIRLPPVQROKIQSATLEEBEPSLSDEASKPLNETLIDLQ----- 124

QY 122 IIHTSRKSPSPAGSPAGSINSFYVYCKGQCRVQSGKRLRVQSGTCRQATLTTLTQGP 181

DB 125 -LESERLNTDTEERVRAKA--HFVHC-SQCDKLCKGKLRVRCALCKGAFVHRDPE 179

QY 182 CWDVLIIPRMGECQSPH--CPGTS-----AEFFPKCGAHPRT-SDKETPVALHLIATN 232

DB 180 CWDVLIIPRMGECQSPH--CPGTS-----AEFFPKCGAHPRT-SDKETPVALHLIATN 239

QY 233 SRNTCTCTDVRSPLVLFQCNRRHVICLDCEHLYCVTRLNDQFVHDPOLGYSILPCVAG 292

DB 240 IKRVPCFCLCTDVRSPLVLFQCNRRHVICLDCEHLYCVTRLNDQFVHDPOLGYSILPCVAG 299

QY 293 CPNSLILKEIHFRILGEGQVNRVQYGAEECTLQWGYLCPRPFGAGLLEPPDRKTC 352

DB 300 CEHSFTIEIHHKPLTLREEDYDQRFATVEEYVLQAGVLCPPGQWGLVPEPDRKTC 359

QY 353 ECGNGLGGCFAPACRECKEAYHGEBC-SAVFASGTTTQAYRVDERAARAEWASKETI 411

DB 360 QN---GGGVFCRNCLOGYHIGBCLPBGTSATNSCEYTVDPNRABAEWASVNTI 415

QY 412 KTTYPCPCRHVPVAKNGGCMHMKCPQPCRLWCNCGCEMNRCVGMGDHWF 463

DB 416 KVTSPCKPCRTPTERDGCWHMVCTTRAGCGFEWCWVQTEWTRDCWGAHWF 467

RESULT 15

ABB67519
ID AAB67519 standard; protein; 250 AA.

AC AAB67519;

DT 29-MAY-2001 (first entry)

DE Amino acid sequence of a murine parkin2 polypeptide.

KM Parkinson's disease; 6q25.2-27; neurodegenerative disease;

KM Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;

KM Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;

KM brain tumour; head trauma; stroke; vascular irregularity;

KM metabolic irregularity.

OS Mus sp.

PN EP1081225-A1.

PR 30-AUG-1999; 99EP-00116766.

PA (BIOF-) BIOFRONTIERA PHARM GMBH.

PI Luebbert H;

DR WPI; 2001-212797/22.

XX N-PSTDB; AAF55246.

PT New polynucleotides encoding mouse parkin2 protein, useful for producing a transgenic non-human animal as an animal model for neurodegenerative diseases.

PS Claim 7; Page 21-22; 62pp; English.

CC The present sequence represents a murine parkin2 polypeptide. Mutations or deletions in the parkin2 gene cause Parkinson's disease in humans. The human parkin2 gene is located in gene region 6q25.2-27. Parkin2 CC polypeptides and polynucleotides are useful for analysing CC neurodegenerative diseases. They are also useful for testing the efficacy of the treatment of a neurodegenerative disease such as Parkinson's disease, Alzheimer's disease, Huntington's disease, amyotrophic lateral

CC scleriosis, Multi-system atrophy, Wilson's disease, Prion
CC disease, and secondary causes inducing Parkinson's syndromes like toxins,
CC drugs, brain tumours, head trauma, stroke, vascular irregularities or
CC metabolic irregularities, associated with a less active or non-active
CC parkin protein
XX

SQ Sequence 250 AA;

Query Match 38.6%; Score 1002; DB 4; Length 250;

Best Local Similarity 77.6%; Pred. No. 1,7e-79; Mismatches 0; Gaps 0;

Matches 190; Conservative 20; Indels 0; Gaps 0;

| | | | | | | | | | | |
|----|-----|---------------------|--------|-------|-------------|---------|----------------|-------|----------------------|----------------------|
| Qy | 1 | MIVFVRFNSSHGPFVEVDS | DTSI | FQ | LKEVYAKQGV | PADQLRV | IFAGKELRNDWTVQ | ND | 60 | |
| Db | 1 | MIVFVRFNSSYGFPVEVDS | DTSI | LQ | LKEVYAKRQGV | PADQLRV | IFAGKELPNHLTVQ | ND | 60 | |
| Qy | 61 | LDQOSIVHIVQRPWRKQGM | ENATG | DDP | RNMAAGC | EREPOS | LTRVPLSSVLP | GD | SVGLA 120 | |
| Db | 61 | LEQOSIVHIVQRPWRKQGM | ENATG | DDP | RNMAAGC | EREPOS | LTRVPLSSVLP | GD | SVGLA 120 | |
| Qy | 121 | VILHTDSRKDSPPAGSP | AGSI | YNS | FFVY | CKGP | CORVOPGKLR | VQ | CGSTGROATLTLTQGP 180 | |
| Db | 121 | VILDTDSRKDSZAAKGP | AVKPT | YNS | FFIY | CKGP | CHKVOPGKLR | VQ | CGCTCKQATLTLAQP 180 | |
| Qy | 181 | SCMDVLI | PNRMSG | ECQSP | HCP | ETSA | AEFFKCGA | HPTSD | KETPV | ALHLIATNSRNITCIT 240 |
| Db | 181 | SCMDVLI | PNRMSG | ECQSP | DCP | ETRA | AEFFKCGA | HPTSD | KETPV | ALNLITSNRISIPCIA 240 |
| Qy | 241 | CTDVR | 245 | | | | | | | |
| Db | 241 | CTDVR | 245 | | | | | | | |

Search completed: February 23, 2005, 13:50:13
Job time : 103.258 secs

This Page Blank (uspro)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:36:20 / Search time 25.8801 Seconds
(without alignments)
1341.256 Million cell updates/sec

Title: US-10-622-817-9

Perfect score: 2596
Sequence: 1 MIVFVRFNSSHGFPEVDS...CNNGCEMNRVCMDHWPDV 465

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/5C_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/5D_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----------------------|----------------------|
| 1 | 2596 | 100.0 | 465 | US-09-601-844B-2 | Sequence 2, Appl 1 |
| 2 | 2596 | 100.0 | 465 | US-09-949-016-6516 | Sequence 6516, Ap |
| 3 | 2405 | 92.6 | 437 | US-09-601-844B-4 | Sequence 4, Appl 1 |
| 4 | 2405 | 92.6 | 437 | US-09-601-844B-4 | Sequence 6853, Ap |
| 5 | 1696.5 | 65.4 | 316 | US-09-949-016-6854 | Sequence 6854, Ap |
| 6 | 370.5 | 14.3 | 117 | US-09-270-767-32183 | Sequence 32183, A |
| 7 | 240.5 | 9.3 | 503 | US-09-914-259-13 | Sequence 13, Appl 1 |
| 8 | 222 | 8.6 | 474 | US-09-354-221-2 | Sequence 2, Appl 1 |
| 9 | 222 | 8.6 | 487 | US-09-949-016-10111 | Sequence 10111, A |
| 10 | 191 | 7.4 | 506 | US-09-248-796A-15410 | Sequence 15410, A |
| 11 | 167 | 6.4 | 587 | US-08-398-008A-2 | Sequence 2, Appl 1 |
| 12 | 167 | 6.4 | 587 | US-08-893-333-2 | Sequence 2, Appl 1 |
| 13 | 149 | 5.7 | 464 | US-09-538-092-598 | Sequence 598, Appl 1 |
| 14 | 141 | 5.4 | 328 | US-09-248-796A-16471 | Sequence 16471, A |
| 15 | 133 | 5.1 | 352 | US-08-854-764-2 | Sequence 2, Appl 1 |
| 16 | 133 | 5.1 | 352 | PCT-US95-09377-2 | Sequence 2, Appl 1 |
| 17 | 131.5 | 5.1 | 1106 | US-09-949-016-9626 | Sequence 9626, Ap |
| 18 | 130.5 | 5.0 | 1242 | US-09-488-270A-2 | Sequence 2, Appl 1 |
| 19 | 128.5 | 4.9 | 229 | US-08-726-306A-23 | Sequence 23, Appl 1 |
| 20 | 128.5 | 4.9 | 229 | US-08-840-146-20 | Sequence 20, Appl 1 |
| 21 | 128.5 | 4.9 | 229 | US-09-360-220-20 | Sequence 20, Appl 1 |
| 22 | 128 | 4.9 | 160 | US-09-370-838-205 | Sequence 205, App |
| 23 | 128 | 4.9 | 160 | US-09-854-133-205 | Sequence 205, App |
| 24 | 127 | 4.9 | 2321 | US-09-230-652-2 | Sequence 2, Appl 1 |
| 25 | 126 | 4.9 | 237 | US-09-248-796A-19062 | Sequence 19062, A |
| 26 | 126 | 4.9 | 846 | US-08-149-103-3 | Sequence 3, Appl 1 |
| 27 | 126 | 4.9 | 846 | US-08-451-883-3 | Sequence 3, Appl 1 |

| | | | | | | |
|----|-------|-----|------|---|----------------------|---------------------|
| 28 | 126 | 4.9 | 873 | 1 | US-08-393-734-2 | Sequence 2, Appl 1 |
| 29 | 126 | 4.9 | 873 | 3 | US-08-894-489-2 | Sequence 2, Appl 1 |
| 30 | 126 | 4.9 | 904 | 4 | US-09-949-016-9528 | Sequence 9528, Ap |
| 31 | 125.5 | 4.8 | 103 | 2 | US-08-771-201-9 | Sequence 9, Appl 1 |
| 32 | 125 | 4.8 | 76 | 1 | US-08-232-815-2 | Sequence 2, Appl 1 |
| 33 | 125 | 4.8 | 76 | 1 | US-08-350-906-2 | Sequence 2, Appl 1 |
| 34 | 125 | 4.8 | 76 | 1 | US-09-601-844B-61 | Sequence 61, Appl 1 |
| 35 | 125 | 4.8 | 76 | 5 | PCT-US95-04536-2 | Sequence 2, Appl 1 |
| 36 | 125 | 4.8 | 206 | 4 | US-09-248-796A-19430 | Sequence 19430, A |
| 37 | 125 | 4.8 | 1121 | 1 | US-07-789-915A-2 | Sequence 2, Appl 1 |
| 38 | 125 | 4.8 | 1121 | 1 | US-08-005-002C-2 | Sequence 2, Appl 1 |
| 39 | 125 | 4.8 | 1121 | 1 | US-08-487-203A-2 | Sequence 2, Appl 1 |
| 40 | 124.5 | 4.8 | 77 | 6 | 5510474-3 | Patent No. 5510474 |
| 41 | 124.5 | 4.8 | 77 | 6 | 5510474-3 | Patent No. 5510474 |
| 42 | 124.5 | 4.8 | 305 | 4 | US-09-866-153-6 | Sequence 6, Appl 1 |
| 43 | 124.5 | 4.8 | 305 | 4 | US-09-693-467A-6 | Sequence 6, Appl 1 |
| 44 | 124.5 | 4.8 | 305 | 4 | US-09-270-976-6 | Sequence 6, Appl 1 |
| 45 | 124.5 | 4.8 | 381 | 4 | US-09-866-153-9 | Sequence 9, Appl 1 |

ALIGNMENTS

| | | | | | | |
|--|-----|---------------------|----------|--------------|---------|------------------|
| RESULT 1 | | | | | | |
| US-09-601-844B-2 | | | | | | |
| Sequence 2, Application US/09601844B | | | | | | |
| Patent No. 6716621 | | | | | | |
| GENERAL INFORMATION: | | | | | | |
| APPLICANT: Shimizu, No. 6716621yoshi | | | | | | |
| APPLICANT: Mizuno, Yoshikuni | | | | | | |
| TITLE OF INVENTION: Isolated DNA or Gene Responsible for Parkinson's Disease | | | | | | |
| FILE REFERENCE: 0652.211000 | | | | | | |
| CURRENT APPLICATION NUMBER: US/09/601,844B | | | | | | |
| PRIOR FILING DATE: 2000-08-09 | | | | | | |
| PRIOR APPLICATION NUMBER: PCT/JP99/00545 | | | | | | |
| PRIOR FILING DATE: 1999-02-09 | | | | | | |
| NUMBER OF SEQ ID NOS: 70 | | | | | | |
| SOFTWARE: PatentIn version 3.1 | | | | | | |
| SEQ ID NO 2 | | | | | | |
| LENGTH: 465 | | | | | | |
| TYPE: PRT | | | | | | |
| ORGANISM: Homo sapiens | | | | | | |
| US-09-601-844B-2 | | | | | | |
| Query Match | | | | | | |
| Best Local Similarity 100.0%; Score 2596; DB 4; Length 465; | | | | | | |
| Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | | | |
| QY | 1 | MIVFVRFNSSHGFPEVDS | DTSI | FOLKEVVARQGV | PADQLRV | IFAGKELRNDMTVQNC |
| DB | 1 | MIVFVRFNSSHGFPEVDS | DTSI | FOLKEVVARQGV | PADQLRV | IFAGKELRNDMTVQNC |
| QY | 61 | IDQOSIVHIVQRPWKGQEM | ATGDDPR | NAAGCEREPQ | SLTRVDL | SSSVLP |
| DB | 61 | IDQOSIVHIVQRPWKGQEM | ATGDDPR | NAAGCEREPQ | SLTRVDL | SSSVLP |
| QY | 121 | VILHDSRVDSP | PASPAR | STVNSFYVY | CKPQ | RVQSGKLRVQ |
| DB | 121 | VILHDSRVDSP | PASPAR | STVNSFYVY | CKPQ | RVQSGKLRVQ |
| QY | 121 | VILHDSRVDSP | PASPAR | STVNSFYVY | CKPQ | RVQSGKLRVQ |
| DB | 121 | VILHDSRVDSP | PASPAR | STVNSFYVY | CKPQ | RVQSGKLRVQ |
| QY | 181 | SCMDVLLIPNRSG | QSPHCP | GTSAEP | FKCAHPT | SKETFPV |
| DB | 181 | SCMDVLLIPNRSG | QSPHCP | GTSAEP | FKCAHPT | SKETFPV |
| QY | 241 | CTDVASPVLV | FOGNSRHV | ICLDC | FLHYCV | TRLNDROFV |
| DB | 241 | CTDVASPVLV | FOGNSRHV | ICLDC | FLHYCV | TRLNDROFV |
| QY | 301 | LHFRFLLIGEO | VNRVQ | QVGAEC | VCVOM | GVLCPR |
| DB | 301 | LHFRFLLIGEO | VNRVQ | QVGAEC | VCVOM | GVLCPR |
| QY | 301 | LHFRFLLIGEO | VNRVQ | QVGAEC | VCVOM | GVLCPR |
| DB | 301 | LHFRFLLIGEO | VNRVQ | QVGAEC | VCVOM | GVLCPR |
| QY | 361 | GPAFCRECKE | AVHEB | CSAVF | EASGTT | QAVYVDR |
| DB | 361 | GPAFCRECKE | AVHEB | CSAVF | EASGTT | QAVYVDR |

```
Db      361  GFACRCEKAYHGECSAVFEASGTTQAYRDEBAAOAWEAASKETIKTTKPCR 420
Qy      421  CHVEKNGGCMHKCPQPCRLIEMCNCGEMNRVCMGDHMFV 465
Db      421  CHVEKNGGCMHKCPQPCRLIEMCNCGEMNRVCMGDHMFV 465
```

RESULT 2

```
US-09-949-016-6516
Sequence 6516, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6516
LENGTH: 465
TYPE: PRT
ORGANISM: Human
US-09-949-016-6516
```

```
Query Match      100.0%; Score 2596; DB 4; Length 465;
Best Local Similarity 100.0%; Pred. No. 8.4e-251;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MIVFVRFNSSHGPFVEVDSDTSIFOLKEVAKRGVPADQLRVIFAGKELRNDWTYQNC 60
Db      1  MIVFVRFNSSHGPFVEVDSDTSIFOLKEVAKRGVPADQLRVIFAGKELRNDWTYQNC 60

Qy      61  LDQOSIVHIVORPMKRGEMNATGDDPRNAAGCGEREPQSLTRVDLSSTVLPDGSVGLA 120
Db      61  LDQOSIVHIVORPMKRGEMNATGDDPRNAAGCGEREPQSLTRVDLSSTVLPDGSVGLA 120

Qy      121  VILHTDSRKDSPAGSPAGRSYNSFYVCKGRCQVPGKLRVQCSYCRQATLTLLTQGP 180
Db      121  VILHTDSRKDSPAGSPAGRSYNSFYVCKGRCQVPGKLRVQCSYCRQATLTLLTQGP 180

Qy      181  SCWDVLIIPNRMSGECQSPHCPGTSAEFFPKGAHPTSDKETPVALLIATNSNITCIT 240
Db      181  SCWDVLIIPNRMSGECQSPHCPGTSAEFFPKGAHPTSDKETPVALLIATNSNITCIT 240

Qy      241  CTDVRSFVLVFOCNSRHVYICLDCPHLYCYTRILNDRQFVHPDQGLYSIPVAGCPSNL 300
Db      241  CTDVRSFVLVFOCNSRHVYICLDCPHLYCYTRILNDRQFVHPDQGLYSIPVAGCPSNL 300

Qy      301  LHHFRILGEBQYNNRYOQYGAEECVLQMGVLCPRPGCAGLLEPPDQRYTCGGNGLGC 360
Db      301  LHHFRILGEBQYNNRYOQYGAEECVLQMGVLCPRPGCAGLLEPPDQRYTCGGNGLGC 360

Qy      361  GFACRCEKAYHGECSAVFEASGTTQAYRDEBAAOAWEAASKETIKTTKPCR 420
Db      361  GFACRCEKAYHGECSAVFEASGTTQAYRDEBAAOAWEAASKETIKTTKPCR 420

Qy      421  CHVEKNGGCMHKCPQPCRLIEMCNCGEMNRVCMGDHMFV 465
Db      421  CHVEKNGGCMHKCPQPCRLIEMCNCGEMNRVCMGDHMFV 465
```

RESULT 3

```
US-09-601-844B-4
Sequence 4, Application US/09601844B
```

```
Patent No. 6716621
GENERAL INFORMATION:
APPLICANT: Shimizu, No. 6716621uyoshi
APPLICANT: Mizuno, Yoshikuni
TITLE OF INVENTION: Isolated DNA or Gene Responsible for Parkinson's Disease
FILE REFERENCE: 0652.211000
CURRENT APPLICATION NUMBER: US/09/601,844B
PRIOR FILING DATE: 2000-08-09
PRIOR APPLICATION NUMBER: PCT/JP99/00545
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 437
TYPE: PRT
ORGANISM: Homo sapiens
US-09-601-844B-4
```

```
Query Match      92.6%; Score 2405; DB 4; Length 437;
Best Local Similarity 94.0%; Pred. No. 9.6e-232;
Matches 437; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

Qy      1  MIVFVRFNSSHGPFVEVDSDTSIFOLKEVAKRGVPADQLRVIFAGKELRNDWTYQNC 60
Db      1  MIVFVRFNSSHGPFVEVDSDTSIFOLKEVAKRGVPADQLRVIFAGKELRNDWTYQNC 60

Qy      61  LDQOSIVHIVORPMKRGEMNATGDDPRNAAGCGEREPQSLTRVDLSSTVLPDGSVGLA 120
Db      61  LDQOSIVHIVORPMKRGEMNATGDDPRNAAGCGEREPQSLTRVDLSSTVLPDGSVGLA 120

Qy      121  VILHTDSRKDSPAGSPAGRSYNSFYVCKGRCQVPGKLRVQCSYCRQATLTLLTQGP 180
Db      121  VILHTDSRKDSPAGSPAGRSYNSFYVCKGRCQVPGKLRVQCSYCRQATLTLLTQGP 180

Qy      181  SCWDVLIIPNRMSGECQSPHCPGTSAEFFPKGAHPTSDKETPVALLIATNSNITCIT 240
Db      179  -----BFFPKGAHPTSDKETPVALLIATNSNITCIT 212

Qy      241  CTDVRSFVLVFOCNSRHVYICLDCPHLYCYTRILNDRQFVHPDQGLYSIPVAGCPSNL 300
Db      213  CTDVRSFVLVFOCNSRHVYICLDCPHLYCYTRILNDRQFVHPDQGLYSIPVAGCPSNL 300

Qy      301  LHHFRILGEBQYNNRYOQYGAEECVLQMGVLCPRPGCAGLLEPPDQRYTCGGNGLGC 360
Db      273  LHHFRILGEBQYNNRYOQYGAEECVLQMGVLCPRPGCAGLLEPPDQRYTCGGNGLGC 360

Qy      361  GFACRCEKAYHGECSAVFEASGTTQAYRDEBAAOAWEAASKETIKTTKPCR 420
Db      333  GFACRCEKAYHGECSAVFEASGTTQAYRDEBAAOAWEAASKETIKTTKPCR 420

Qy      421  CHVEKNGGCMHKCPQPCRLIEMCNCGEMNRVCMGDHMFV 465
Db      393  CHVEKNGGCMHKCPQPCRLIEMCNCGEMNRVCMGDHMFV 437
```

RESULT 4

```
US-09-949-016-6853
Sequence 6853, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
NUMBER OF SEQ ID NOS: 207012
```

SOFTWARE: FaastSeq for windows Version 4.0
 ; SEQ ID NO 6853
 ; LENGTH: 437
 ; TYPE: prt
 ; ORGANISM: Human
 US-09-949-016-6853

Query Match 92.6%; Score 2405; DB 4; Length 437;
 Best Local Similarity 94.0%; Pred. No. 9, 6e-232;
 Matches 437; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

```

QY 1 MIVFVRFNSSHGFPVEVDSITSIFOLKEVYAKRGVPADQLRVIFAGKELRNDMTYQNC 60
DB 1 MIVFVRFNSSHGFPVEVDSITSIFOLKEVYAKRGVPADQLRVIFAGKELRNDMTYQNC 60
QY 61 LDOQSI VHVQRPWRKGGEMNATGDDPRNNAAGCEREPQSLTRVDLSSVLPBGDSVGLA 120
DB 61 LDOQSI VHVQRPWRKGGEMNATGDDPRNNAAGCEREPQSLTRVDLSSVLPBGDSVGLA 120
QY 121 VILHTDSRKDSPAGSPAGRSIYNSFYVCKGPCQVRQPKLRVQCSTGQATLTLLTQGP 180
DB 121 VILHTDSRKDSPAGSPAGRSIYNSFYVCKGPCQVRQPKLRVQCSTGQATLTLLTQGP 180
QY 121 VILHTDSRKDSPAGSPAGRSIYNSFYVCKGPCQVRQPKLRVQCSTGQATLTLLTQGP 178
DB 121 VILHTDSRKDSPAGSPAGRSIYNSFYVCKGPCQVRQPKLRVQCSTGQATLTLLTQGP 178
QY 191 SCMDVLI.PNRMSGCQSPHCPTSAEPFKCGAHPDTSKETPVLAHL.IATNSRNITCIT 240
DB 179 -----BPFKCGAHPDTSKETPVLAHL.IATNSRNITCIT 212
QY 241 CTDVSPVTVFQCNRRHYCLDCCFHL.YCTTRLNDROFVNDPOLGYSLPCVACCPNSLIE 300
DB 213 CTDVSPVTVFQCNRRHYCLDCCFHL.YCTTRLNDROFVNDPOLGYSLPCVACCPNSLIE 272
QY 301 LHHFRILGEEQYNYQYGAEECVLQMGVLCPRPCGAGLLPEPDQRKVTCEGGNGLGC 360
DB 273 LHHFRILGEEQYNYQYGAEECVLQMGVLCPRPCGAGLLPEPDQRKVTCEGGNGLGC 332
QY 361 GFAFCRECKEAYHEBSCAVFEASGTTTQAYRVDERAAQARWEAASKETIKTKPCPR 420
DB 333 GFAFCRECKEAYHEBSCAVFEASGTTTQAYRVDERAAQARWEAASKETIKTKPCPR 392
QY 421 CHVPEKNGGCMHMKCPQPCRLIEWCNCGCEWNRVCMGDHMFV 465
DB 393 CHVPEKNGGCMHMKCPQPCRLIEWCNCGCEWNRVCMGDHMFV 437

```

RESULT 5
 US-09-949-016-6854
 ; Sequence 6854, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FaastSeq for windows Version 4.0
 ; SEQ ID NO 6854
 ; LENGTH: 316
 ; TYPE: prt
 ; ORGANISM: Human
 US-09-949-016-6854

Query Match 65.4%; Score 1696.5; DB 4; Length 316;
 Best Local Similarity 68.0%; Pred. No. 4, 2e-161;
 Matches 316; Conservative 0; Mismatches 0; Indels 149; Gaps 1;

```

QY 1 MIVFVRFNSSHGFPVEVDSITSIFOLKEVYAKRGVPADQLRVIFAGKELRNDMTYQNC 60
DB 1 MIVFVRFNSSHGFPVEVDSITSIFOLKEVYAKRGVPADQLRVIFAGKELRNDMTYQNC 57
QY 61 LDOQSI VHVQRPWRKGGEMNATGDDPRNNAAGCEREPQSLTRVDLSSVLPBGDSVGLA 120
DB 58 ----- 57
QY 121 VILHTDSRKDSPAGSPAGRSIYNSFYVCKGPCQVRQPKLRVQCSTGQATLTLLTQGP 180
DB 58 ----- 57
QY 181 SCMDVLI.PNRMSGCQSPHCPTSAEPFKCGAHPDTSKETPVLAHL.IATNSRNITCIT 240
DB 58 ----- 57
QY 241 CTDVSPVTVFQCNRRHYCLDCCFHL.YCTTRLNDROFVNDPOLGYSLPCVACCPNSLIE 300
DB 92 CTDVSPVTVFQCNRRHYCLDCCFHL.YCTTRLNDROFVNDPOLGYSLPCVACCPNSLIE 151
QY 301 LHHFRILGEEQYNYQYGAEECVLQMGVLCPRPCGAGLLPEPDQRKVTCEGGNGLGC 360
DB 152 LHHFRILGEEQYNYQYGAEECVLQMGVLCPRPCGAGLLPEPDQRKVTCEGGNGLGC 211
QY 361 GFAFCRECKEAYHEBSCAVFEASGTTTQAYRVDERAAQARWEAASKETIKTKPCPR 420
DB 212 GFAFCRECKEAYHEBSCAVFEASGTTTQAYRVDERAAQARWEAASKETIKTKPCPR 271
QY 421 CHVPEKNGGCMHMKCPQPCRLIEWCNCGCEWNRVCMGDHMFV 465
DB 272 CHVPEKNGGCMHMKCPQPCRLIEWCNCGCEWNRVCMGDHMFV 316

```

RESULT 6
 US-09-270-767-32183
 ; Sequence 32183, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 32183
 ; LENGTH: 117
 ; TYPE: prt
 ; ORGANISM: Drosophila melanogaster
 US-09-270-767-32183

Query Match 14.3%; Score 370.5; DB 4; Length 117;
 Best Local Similarity 53.3%; Pred. No. 3, 8e-29;
 Matches 64; Conservative 11; Mismatches 40; Indels 5; Gaps 2;

```

QY 345 PDQRVTCGGNGLGGFAGFCECKEAYHEGEC-SAVFEASGTTTQAYRVDERAAQARW 403
DB 1 PDQRVTCGN---CGGVFCRNCLOGHIGCLBEGGASATNSCEYVDPNRAASARW 56
QY 404 EAASKETIKTKPCPRCHVPEKNGGCMHMKCPQPCRLIEWCNCGCEWNRVCMGDHMF 463
DB 57 DEASNVITIKVTKPCPKRTPTERRDGGCMHMKCTAAGCGFEMCWQCQTEWTDGCAHWF 116

```

RESULT 7
 US-09-914-259-13
 ; Sequence 13, Application US/09914259
 ; Patent No. 6495336
 ; GENERAL INFORMATION:
 ; APPLICANT: Nakowski, Lee
 ; APPLICANT: Hyman, Paul
 ; APPLICANT: Williams, Mark
 ; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES

FILE REFERENCE: 8471-010-999
CURRENT APPLICATION NUMBER: US/09/914,259
CURRENT FILING DATE: 2000-11-21
NUMBER OF SEQ ID NOS: 180
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13
LENGTH: 503
TYPE: PRF
ORGANISM: Drosophila melanogaster
US-09-914-259-13

Query Match
Best Local Similarity 27.5%; Score 240.5; DB 4; Length 503;
Matches 72; Conservative 42; Mismatches 111; Indels 37; Gaps 13;

QY 209 FFKCAHPTSDKETPVALLHATNSRNITCTCTDVASPLVYQCSNRVITCLDCEPHLYC 268
DB 105 FFKC-AHYINPNAITKOKTSQCEBEICFSQLPDSMAGLECGHRCFPCMCHEYL 163
QY 269 VTRLNDQFVNDPDLGSLPCVA-GCPNSLKEHFRILGEOYV-RYQYGAECVLO 326
DB 164 STK-----IVASLGOTISCAHGC-DILVDVTVANLVTDAVRVYKQULITNSFY-E 215
QY 327 MGVLT-CPRPGCG-AGLLEPDRKVTCEGNGJGCGFAFCECKEAYHEG-EGSAVFE 382
DB 216 CNDLIRWCPESVDCYAVKVPYAEPRVHCK-----CGHVECFACGENMHPVXCRWL- 267
QY 383 ASGTTQAVYVDEBRAEQARWEASKEITIKTTKPCPRCHVPEKNGGCMKCPQPCQR 442
DB 268 -----KKWIKKCCDDSETSNMTAAN-----TKCPCRCSTVIEEDGCHNMVCKNQCK 315
QY 443 LEKWNCGCEMNRVCWGDHMF 464
DB 316 NEFCWVCLGSMER--HGSSWN 335

RESULT 8
US-09-354-221-2
Sequence 2, Application US/09354221
Patent No. 6699714
GENERAL INFORMATION:
APPLICANT: Chang, Chawmshang
TITLE OF INVENTION: Androgen Receptor Coactivators
FILE REFERENCE: 920920.90011
CURRENT APPLICATION NUMBER: US/09/354,221
CURRENT FILING DATE: 1999-07-15
EARLIER APPLICATION NUMBER: US 60/100,243
EARLIER FILING DATE: 1998-09-14
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 474
TYPE: PRF
ORGANISM: Homo sapien
US-09-354-221-2

Query Match
Best Local Similarity 8.6%; Score 222; DB 4; Length 474;
Matches 71; Conservative 27; Mismatches 97; Indels 60; Gaps 13;

QY 232 NSRNITC--ITCTDVASPLVYF-QCNSRHVITCLDCEPHLYCVTRLNDQFVNDPDLGSLP 288
DB 214 NSKTLFLCSICFCEKJGSECMYFLEC--RHVYCKAQLDYEIQIRDOQ-----VQ 261
QY 289 CVAGCPN-----SLIKELHFRILGEOYVRYQYGAECVLOMGVLT-CPRPGCG 338
DB 262 CL-NCPEKPCSVATPGQVKL-----VEALFARYDRLLQSLDLMADVVCPRPCQ 315
QY 339 AGLLEPDRKVTCEGNGJGCGFAFCECKEAYH-EGESAVFEASGTTQAY----- 391
DB 316 LPVWDEP-----CTMGICSSCNFAFCTLCRLTYHGVSPCKYTAELMDLRNEYILOADEA 370
QY 392 -----RVDERAAEQARWEASKEITIKTTKPCPRCHVPEKNGGCMKCPQPCRL 444

DB 371 NKRLLDQRYGKRVIOKALEMESKEWLEKNSKSCPCGTPRIKLDGCKNMTG--TGCWQY 428
QY 445 WCNWNGCEMNRVCWG 459
DB 429 FCW-----ICWG 435

RESULT 9
US-09-949-016-10111
Sequence 10111, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10111
LENGTH: 487
TYPE: PRF
ORGANISM: Human
US-09-949-016-10111

Query Match
Best Local Similarity 8.6%; Score 222; DB 4; Length 487;
Matches 71; Conservative 27; Mismatches 97; Indels 60; Gaps 13;

QY 232 NSRNITC--ITCTDVASPLVYF-QCNSRHVITCLDCEPHLYCVTRLNDQFVNDPDLGSLP 288
DB 227 NSKTLFLCSICFCEKJGSECMYFLEC--RHVYCKAQLDYEIQIRDOQ-----VQ 274
QY 289 CVAGCPN-----SLIKELHFRILGEOYVRYQYGAECVLOMGVLT-CPRPGCG 338
DB 275 CL-NCPEKPCSVATPGQVKL-----VEALFARYDRLLQSLDLMADVVCPRPCQ 328
QY 339 AGLLEPDRKVTCEGNGJGCGFAFCECKEAYH-EGESAVFEASGTTQAY----- 391
DB 329 LPVWDEP-----CTMGICSSCNFAFCTLCRLTYHGVSPCKYTAELMDLRNEYILOADEA 383
QY 392 -----RVDERAAEQARWEASKEITIKTTKPCPRCHVPEKNGGCMKCPQPCRL 444
DB 384 NKRLLDQRYGKRVIOKALEMESKEWLEKNSKSCPCGTPRIKLDGCKNMTG--TGCWQY 441
QY 445 WCNWNGCEMNRVCWG 459
DB 442 FCW-----ICWG 448

RESULT 10
US-09-248-796A-15410
Sequence 15410, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:57:26 ; Search time 73.0113 Seconds
(without alignments)
2084.158 Million cell updates/sec

Title: US-10-622-817-9
Perfect score: 2596
Sequence: 1 MIVFVRPNSSHGFVEVDSD.....CNNCGCEMNRRVCMGDHMFV 465

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
```

```
Database : Published Applications AA:*
1: /cgn2_6/ptodataa/1/pubppaa/US07_PUBCOMB.pcp:*
2: /cgn2_6/ptodataa/1/pubppaa/PCT_NEW_PUB.pcp:*
3: /cgn2_6/ptodataa/1/pubppaa/US06_NEW_PUB.pcp:*
4: /cgn2_6/ptodataa/1/pubppaa/US06_PUBCOMB.pcp:*
5: /cgn2_6/ptodataa/1/pubppaa/US07_NEW_PUB.pcp:*
6: /cgn2_6/ptodataa/1/pubppaa/PCTUS_PUBCOMB.pcp:*
7: /cgn2_6/ptodataa/1/pubppaa/US08_NEW_PUB.pcp:*
8: /cgn2_6/ptodataa/1/pubppaa/US08_PUBCOMB.pcp:*
9: /cgn2_6/ptodataa/1/pubppaa/US09A_PUBCOMB.pcp:*
10: /cgn2_6/ptodataa/1/pubppaa/US09B_PUBCOMB.pcp:*
11: /cgn2_6/ptodataa/1/pubppaa/US09C_PUBCOMB.pcp:*
12: /cgn2_6/ptodataa/1/pubppaa/US09_NEW_PUB.pcp:*
13: /cgn2_6/ptodataa/1/pubppaa/US10A_PUBCOMB.pcp:*
14: /cgn2_6/ptodataa/1/pubppaa/US10B_PUBCOMB.pcp:*
15: /cgn2_6/ptodataa/1/pubppaa/US10C_PUBCOMB.pcp:*
16: /cgn2_6/ptodataa/1/pubppaa/US10D_PUBCOMB.pcp:*
17: /cgn2_6/ptodataa/1/pubppaa/US10_NEW_PUB.pcp:*
18: /cgn2_6/ptodataa/1/pubppaa/US11_NEW_PUB.pcp:*
19: /cgn2_6/ptodataa/1/pubppaa/US00_NEW_PUB.pcp:*
20: /cgn2_6/ptodataa/1/pubppaa/US60_PUBCOMB.pcp:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------------------|--------------------|
| 1 | 2596 | 100.0 | 465 | 16 | US-10-473-226-2 | Sequence 2, Appl1 |
| 2 | 2596 | 100.0 | 465 | 16 | US-10-776-604-2 | Sequence 2, Appl1 |
| 3 | 2596 | 100.0 | 465 | 17 | US-10-839-688-9 | Sequence 9, Appl1 |
| 4 | 2584 | 99.5 | 465 | 16 | US-10-473-226-4 | Sequence 4, Appl1 |
| 5 | 2405 | 92.6 | 437 | 16 | US-10-776-604-4 | Sequence 4, Appl1 |
| 6 | 870 | 33.5 | 156 | 9 | US-09-785-548-4 | Sequence 4, Appl1 |
| 7 | 544.5 | 21.0 | 386 | 14 | US-10-239-249-2 | Sequence 2, Appl1 |
| 8 | 373 | 14.4 | 64 | 15 | US-10-313-203-10 | Sequence 10, Appl1 |
| 9 | 323 | 12.4 | 56 | 15 | US-10-313-203-17 | Sequence 17, Appl1 |
| 10 | 263 | 10.1 | 46 | 9 | US-09-864-761-36750 | Sequence 36750, A |
| 11 | 246 | 9.5 | 525 | 16 | US-10-437-963-199787 | Sequence 199787, A |
| 12 | 242.5 | 9.3 | 604 | 15 | US-10-425-111-55407 | Sequence 55407, A |
| 13 | 240.5 | 9.3 | 503 | 14 | US-10-080-608A-13 | Sequence 13, Appl1 |

| | | | | | | |
|----|-------|-----|------|----|----------------------|---------------------|
| 14 | 230.5 | 9.3 | 503 | 15 | US-10-370-685-102 | Sequence 102, Appl |
| 15 | 243.5 | 9.0 | 585 | 15 | US-10-424-599-147673 | Sequence 147673, A |
| 16 | 224.5 | 8.6 | 648 | 15 | US-10-425-114-63379 | Sequence 63379, A |
| 17 | 222 | 8.6 | 474 | 15 | US-09-808-387-42 | Sequence 42, Appl |
| 18 | 222 | 8.6 | 474 | 14 | US-10-442-754-2 | Sequence 2, Appl1 |
| 19 | 222 | 8.6 | 474 | 16 | US-10-408-7655-572 | Sequence 572, Appl |
| 20 | 222 | 8.6 | 477 | 15 | US-10-264-049-3019 | Sequence 3019, Appl |
| 21 | 219.5 | 8.5 | 1753 | 14 | US-10-146-473-44 | Sequence 44, Appl |
| 22 | 219.5 | 8.5 | 1753 | 15 | US-10-276-774-1824 | Sequence 1824, Appl |
| 23 | 215.5 | 8.5 | 2517 | 15 | US-10-313-203-2 | Sequence 2, Appl1 |
| 24 | 215 | 8.3 | 32 | 15 | US-10-313-203-18 | Sequence 18, Appl1 |
| 25 | 207.5 | 8.0 | 623 | 16 | US-10-437-963-148905 | Sequence 148905, A |
| 26 | 206.5 | 8.0 | 452 | 15 | US-10-425-114-38443 | Sequence 38443, A |
| 27 | 205.5 | 7.9 | 449 | 15 | US-10-425-114-39657 | Sequence 39657, A |
| 28 | 204 | 7.9 | 624 | 15 | US-10-374-7804-1727 | Sequence 1727, Appl |
| 29 | 204 | 7.9 | 624 | 16 | US-10-437-963-114793 | Sequence 14793, A |
| 30 | 202.5 | 7.8 | 289 | 16 | US-10-437-963-133385 | Sequence 133385, A |
| 31 | 201 | 7.7 | 333 | 15 | US-10-424-599-217796 | Sequence 217796, A |
| 32 | 200 | 7.7 | 593 | 15 | US-10-325-0656-854 | Sequence 854, Appl |
| 33 | 200 | 7.7 | 593 | 15 | US-10-325-067-56 | Sequence 56, Appl |
| 34 | 200 | 7.7 | 593 | 15 | US-10-374-7804-330 | Sequence 330, Appl |
| 35 | 195.5 | 7.5 | 541 | 15 | US-10-424-599-265215 | Sequence 26515, A |
| 36 | 193 | 7.4 | 606 | 16 | US-10-437-963-113378 | Sequence 113378, A |
| 37 | 192 | 7.4 | 354 | 15 | US-10-425-114-55248 | Sequence 55248, A |
| 38 | 191.5 | 7.4 | 540 | 16 | US-10-437-963-168787 | Sequence 168787, A |
| 39 | 191 | 7.4 | 338 | 15 | US-10-425-114-56467 | Sequence 56467, A |
| 40 | 184 | 7.1 | 583 | 16 | US-10-437-963-124556 | Sequence 124556, A |
| 41 | 182.5 | 7.0 | 522 | 16 | US-10-437-963-184651 | Sequence 184651, A |
| 42 | 180 | 6.9 | 250 | 16 | US-10-437-963-152161 | Sequence 152161, A |
| 43 | 180 | 6.9 | 394 | 15 | US-10-425-114-54725 | Sequence 54725, A |
| 44 | 180 | 6.9 | 557 | 15 | US-10-424-599-212592 | Sequence 212592, A |
| 45 | 177.5 | 6.8 | 396 | 15 | US-10-425-114-40231 | Sequence 40231, A |

ALIGNMENTS

```

RESULT 1
US-10-473-226-2
; Sequence 2, Application US/10473226
; Publication No. US20040198650A1
; GENERAL INFORMATION:
; APPLICANT: NsGene A/S
; TITLE OF INVENTION: Means for inhibiting proteolytical processing of Parkin
; FILE REFERENCE: 506-204-MO
; CURRENT APPLICATION NUMBER: US/10/473,226
; CURRENT FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: DK PA 2001 00525
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: US 60/281,286
; PRIOR FILING DATE: 2001-04-03
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (1) - ( )
; OTHER INFORMATION: Native Parkin
US-10-473-226-2

```

QY 1 MIVFRRSSSGFPEVSDSTSI FOLKEVARQGVPAQQLRVIFAGKELRDMTVQNC 60

Db 1 MIVFRRSSSGFPEVSDSTSI FOLKEVARQGVPAQQLRVIFAGKELRDMTVQNC 60

QY 61 LDQGSIVHIVRRPWKGGEMNATGGDDPPNNAAGGGEREQSLTRVLLSSVLPDGSIGLA 120

```

Db      61 LDQOSIVHIVORPMKGGEMNATGDDPRNAAGCEREPOSTLRVDLSSSVLPGBSVGLA 120
Qy      121 VILHTDSRKDSPSPAGSPAGRSIYNSFYVYCKGPCQORVOPGKLRVQSTCRQATLTLTQGP 180
Db      121 VILHTDSRKDSPSPAGSPAGRSIYNSFYVYCKGPCQORVOPGKLRVQSTCRQATLTLTQGP 180
Qy      181 SCMDVLIIPNMSGECSPHCPSGTSAEFPFKCGAHTSDKETPVALHLIATNSRNTICIT 240
Db      181 SCMDVLIIPNMSGECSPHCPSGTSAEFPFKCGAHTSDKETPVALHLIATNSRNTICIT 240
Qy      241 CTDVRSPLVLFQCNRSRHVITCLDCEPHLYCTVRLNDROFVHDPOLGSLPCVAGCENSLIKE 300
Db      241 CTDVRSPLVLFQCNRSRHVITCLDCEPHLYCTVRLNDROFVHDPOLGSLPCVAGCENSLIKE 300
Qy      301 LHHFRILGEBQYNNRYOQYGAEBECVLOMGVLCPRPGCAGLLPEPDQRKTYCEGNGLGC 360
Db      301 LHHFRILGEBQYNNRYOQYGAEBECVLOMGVLCPRPGCAGLLPEPDQRKTYCEGNGLGC 360
Qy      361 GFAPCECKEAYHEGSCSAVFEASGTTTOAYRVDERAABOARWEAASKETIKKTKPCPR 420
Db      361 GFAPCECKEAYHEGSCSAVFEASGTTTOAYRVDERAABOARWEAASKETIKKTKPCPR 420
Qy      421 CHVPEKNGGCMHMKCPQPCRLEWCWNCGCEMNRYCMGDHMFV 465
Db      421 CHVPEKNGGCMHMKCPQPCRLEWCWNCGCEMNRYCMGDHMFV 465

```

```

RESULT 2
US-10-776-604-2
; Sequence 2, Application US/10776604
; Publication No. US20050003385A1
; GENERAL INFORMATION:
; APPLICANT: Shimizu, Nobuyoshi
; APPLICANT: Miuno, Yoshikuni
; TITLE OF INVENTION: DMS or Genes Participating in Parkinson's Disease
; FILE REFERENCE: 0652.2110001
; CURRENT APPLICATION NUMBER: US/10/776,604
; PRIOR FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 09/601,844
; PRIOR FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: PCT/JP99/00545
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: JP 10/27531
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-776-604-2

```

```

Query Match      100.0%; Score 2596; DB 16; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.2e-220;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MIVFVFNSSHGFPVEVSDTSTIFOLKEVAVAKRGVPADOLRVTFAGKELRNDMTVQNC 60
Db      1 MIVFVFNSSHGFPVEVSDTSTIFOLKEVAVAKRGVPADOLRVTFAGKELRNDMTVQNC 60
Qy      61 LDQOSIVHIVORPMKGGEMNATGDDPRNAAGCEREPOSTLRVDLSSSVLPGBSVGLA 120
Db      61 LDQOSIVHIVORPMKGGEMNATGDDPRNAAGCEREPOSTLRVDLSSSVLPGBSVGLA 120
Qy      121 VILHTDSRKDSPSPAGSPAGRSIYNSFYVYCKGPCQORVOPGKLRVQSTCRQATLTLTQGP 180
Db      121 VILHTDSRKDSPSPAGSPAGRSIYNSFYVYCKGPCQORVOPGKLRVQSTCRQATLTLTQGP 180
Qy      181 SCMDVLIIPNMSGECSPHCPSGTSAEFPFKCGAHTSDKETPVALHLIATNSRNTICIT 240
Db      181 SCMDVLIIPNMSGECSPHCPSGTSAEFPFKCGAHTSDKETPVALHLIATNSRNTICIT 240
Qy      241 CTDVRSPLVLFQCNRSRHVITCLDCEPHLYCTVRLNDROFVHDPOLGSLPCVAGCENSLIKE 300

```

```

Db      241 CTDVRSPLVLFQCNRSRHVITCLDCEPHLYCTVRLNDROFVHDPOLGSLPCVAGCENSLIKE 300
Qy      301 LHHFRILGEBQYNNRYOQYGAEBECVLOMGVLCPRPGCAGLLPEPDQRKTYCEGNGLGC 360
Db      301 LHHFRILGEBQYNNRYOQYGAEBECVLOMGVLCPRPGCAGLLPEPDQRKTYCEGNGLGC 360
Qy      361 GFAPCECKEAYHEGSCSAVFEASGTTTOAYRVDERAABOARWEAASKETIKKTKPCPR 420
Db      361 GFAPCECKEAYHEGSCSAVFEASGTTTOAYRVDERAABOARWEAASKETIKKTKPCPR 420
Qy      421 CHVPEKNGGCMHMKCPQPCRLEWCWNCGCEMNRYCMGDHMFV 465
Db      421 CHVPEKNGGCMHMKCPQPCRLEWCWNCGCEMNRYCMGDHMFV 465

```

```

RESULT 3
US-10-839-688-9
; Sequence 9, Application US/10839688
; Publication No. US20050014173A1
; GENERAL INFORMATION:
; APPLICANT: Farex, Matthew J.
; TITLE OF INVENTION: PARKINSON'S DISEASE MARKERS
; FILE REFERENCE: 07039-448001
; CURRENT APPLICATION NUMBER: US/10/839,688
; PRIOR FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: US 60/468,832
; PRIOR FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-839-688-9

```

```

Query Match      100.0%; Score 2596; DB 17; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.2e-220;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MIVFVFNSSHGFPVEVSDTSTIFOLKEVAVAKRGVPADOLRVTFAGKELRNDMTVQNC 60
Db      1 MIVFVFNSSHGFPVEVSDTSTIFOLKEVAVAKRGVPADOLRVTFAGKELRNDMTVQNC 60
Qy      61 LDQOSIVHIVORPMKGGEMNATGDDPRNAAGCEREPOSTLRVDLSSSVLPGBSVGLA 120
Db      61 LDQOSIVHIVORPMKGGEMNATGDDPRNAAGCEREPOSTLRVDLSSSVLPGBSVGLA 120
Qy      121 VILHTDSRKDSPSPAGSPAGRSIYNSFYVYCKGPCQORVOPGKLRVQSTCRQATLTLTQGP 180
Db      121 VILHTDSRKDSPSPAGSPAGRSIYNSFYVYCKGPCQORVOPGKLRVQSTCRQATLTLTQGP 180
Qy      181 SCMDVLIIPNMSGECSPHCPSGTSAEFPFKCGAHTSDKETPVALHLIATNSRNTICIT 240
Db      181 SCMDVLIIPNMSGECSPHCPSGTSAEFPFKCGAHTSDKETPVALHLIATNSRNTICIT 240
Qy      241 CTDVRSPLVLFQCNRSRHVITCLDCEPHLYCTVRLNDROFVHDPOLGSLPCVAGCENSLIKE 300
Db      241 CTDVRSPLVLFQCNRSRHVITCLDCEPHLYCTVRLNDROFVHDPOLGSLPCVAGCENSLIKE 300
Qy      301 LHHFRILGEBQYNNRYOQYGAEBECVLOMGVLCPRPGCAGLLPEPDQRKTYCEGNGLGC 360
Db      301 LHHFRILGEBQYNNRYOQYGAEBECVLOMGVLCPRPGCAGLLPEPDQRKTYCEGNGLGC 360
Qy      361 GFAPCECKEAYHEGSCSAVFEASGTTTOAYRVDERAABOARWEAASKETIKKTKPCPR 420
Db      361 GFAPCECKEAYHEGSCSAVFEASGTTTOAYRVDERAABOARWEAASKETIKKTKPCPR 420
Qy      421 CHVPEKNGGCMHMKCPQPCRLEWCWNCGCEMNRYCMGDHMFV 465
Db      421 CHVPEKNGGCMHMKCPQPCRLEWCWNCGCEMNRYCMGDHMFV 465

```

```
RESULT 4
US-10-473-226-4
; Sequence 4, Application US/10473226
; Publication No. US20040198650A1
; GENERAL INFORMATION:
; APPLICANT: NcGene A/S
; TITLE OF INVENTION: Means for inhibiting proteolytical processing of Parkin
; FILE REFERENCE: 506-204-WO
; CURRENT APPLICATION NUMBER: US/10/473,226
; PRIOR FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: DK PA 2001 00525
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: US 60/281,286
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (1)..(1)
; OTHER INFORMATION: Parkin with a D126E mutation
US-10-473-226-4

Query Match          99.6%; Score 2584; DB 16; Length 465;
Best Local Similarity 99.6%; Pred. No. 2.5e-219;
Matches 463; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MIVFRFNSSHGFPVEVSDTSIFOLKEVVARQGVADQLRVIFAGKELRNDMTVQNC 60
DB 1 MIVFRFNSSHGFPVEVSDTSIFOLKEVVARQGVADQLRVIFAGKELRNDMTVQNC 60
QY 61 IDQGSIVHIVQRPWRKQGMNATGDDPRNAAAGCCEREPQSLTRVDLSSVLPDGSVGIA 120
DB 61 IDQGSIVHIVQRPWRKQGMNATGDDPRNAAAGCCEREPQSLTRVDLSSVLPDGSVGIA 120
QY 121 VLTHTDSRKDPPASPAGRSTIYNSFYVCKPCQVOPGKURVOCSTCRQATLTLTQGP 180
DB 121 VLTHTDSRKDPPASPAGRSTIYNSFYVCKPCQVOPGKURVOCSTCRQATLTLTQGP 180
QY 121 VLTHTDSRKDPPASPAGRSTIYNSFYVCKPCQVOPGKURVOCSTCRQATLTLTQGP 180
DB 121 VLTHTDSRKDPPASPAGRSTIYNSFYVCKPCQVOPGKURVOCSTCRQATLTLTQGP 180
QY 181 SCMDVLLIPNRMSGECQSPHCPTSAEFPFKCAHPTSDKETPVALLHATNSRNITCIT 240
DB 181 SCMDVLLIPNRMSGECQSPHCPTSAEFPFKCAHPTSDKETPVALLHATNSRNITCIT 240
QY 241 CTDVASPVLPFCNSRHVYCLDPCFHLVYCVTRLNDRQFVHDPQLGYSLPFCVACCPNSLIKE 300
DB 241 CTDVASPVLPFCNSRHVYCLDPCFHLVYCVTRLNDRQFVHDPQLGYSLPFCVACCPNSLIKE 300
QY 301 LHHFRILGEEQYNNRYQOYGAEECVLQMGVLCPRPCCGAGLLPEPDQRVYTCGGNGLGC 360
DB 301 LHHFRILGEEQYNNRYQOYGAEECVLQMGVLCPRPCCGAGLLPEPDQRVYTCGGNGLGC 360
QY 301 LHHFRILGEEQYNNRYQOYGAEECVLQMGVLCPRPCCGAGLLPEPDQRVYTCGGNGLGC 360
DB 301 LHHFRILGEEQYNNRYQOYGAEECVLQMGVLCPRPCCGAGLLPEPDQRVYTCGGNGLGC 360
QY 361 GFAFCRECKEAYHEBECSAVFPAASGTTTQAYVDEBAAQAEWAASKETIKKTPCPR 420
DB 361 GFAFCRECKEAYHEBECSAVFPAASGTTTQAYVDEBAAQAEWAASKETIKKTPCPR 420
QY 421 CHVPEKNGGCMHMKCPQPCRLBWCNCGCEMNRVCMGDHMFV 465
DB 421 CHVPEKNGGCMHMKCPQPCRLBWCNCGCEMNRVCMGDHMFV 465

RESULT 5
US-10-776-604-4
; Sequence 4, Application US/10776604
; Publication No. US20050003385A1
; GENERAL INFORMATION:
; APPLICANT: Shimizu, Nobuyoshi
; APPLICANT: Mizuno, Yoshiaki
; TITLE OF INVENTION: DNAs or Genes Participating in Parkinson's Disease
; FILE REFERENCE: 0652.2110001
; CURRENT APPLICATION NUMBER: US/10/776,604
```

```
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 09/601,844
; PRIOR FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: PCT/JP99/00545
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: JP 10/27531
; PRIOR FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-776-604-4

Query Match          92.6%; Score 2405; DB 16; Length 437;
Best Local Similarity 94.0%; Pred. No. 1.5e-203;
Matches 437; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

QY 1 MIVFRFNSSHGFPVEVSDTSIFOLKEVVARQGVADQLRVIFAGKELRNDMTVQNC 60
DB 1 MIVFRFNSSHGFPVEVSDTSIFOLKEVVARQGVADQLRVIFAGKELRNDMTVQNC 60
QY 61 IDQGSIVHIVQRPWRKQGMNATGDDPRNAAAGCCEREPQSLTRVDLSSVLPDGSVGIA 120
DB 61 IDQGSIVHIVQRPWRKQGMNATGDDPRNAAAGCCEREPQSLTRVDLSSVLPDGSVGIA 120
QY 121 VLTHTDSRKDPPASPAGRSTIYNSFYVCKPCQVOPGKURVOCSTCRQATLTLTQGP 180
DB 121 VLTHTDSRKDPPASPAGRSTIYNSFYVCKPCQVOPGKURVOCSTCRQATLTLTQGP 180
QY 121 VLTHTDSRKDPPASPAGRSTIYNSFYVCKPCQVOPGKURVOCSTCRQATLTLTQGP 180
DB 121 VLTHTDSRKDPPASPAGRSTIYNSFYVCKPCQVOPGKURVOCSTCRQATLTLTQGP 180
QY 181 SCMDVLLIPNRMSGECQSPHCPTSAEFPFKCAHPTSDKETPVALLHATNSRNITCIT 240
DB 181 SCMDVLLIPNRMSGECQSPHCPTSAEFPFKCAHPTSDKETPVALLHATNSRNITCIT 240
QY 179 -----EFPFKCAHPTSDKETPVALLHATNSRNITCIT 212
DB 179 -----EFPFKCAHPTSDKETPVALLHATNSRNITCIT 212
QY 241 CTDVASPVLPFCNSRHVYCLDPCFHLVYCVTRLNDRQFVHDPQLGYSLPFCVACCPNSLIKE 300
DB 241 CTDVASPVLPFCNSRHVYCLDPCFHLVYCVTRLNDRQFVHDPQLGYSLPFCVACCPNSLIKE 300
QY 301 LHHFRILGEEQYNNRYQOYGAEECVLQMGVLCPRPCCGAGLLPEPDQRVYTCGGNGLGC 360
DB 301 LHHFRILGEEQYNNRYQOYGAEECVLQMGVLCPRPCCGAGLLPEPDQRVYTCGGNGLGC 360
QY 301 LHHFRILGEEQYNNRYQOYGAEECVLQMGVLCPRPCCGAGLLPEPDQRVYTCGGNGLGC 360
DB 301 LHHFRILGEEQYNNRYQOYGAEECVLQMGVLCPRPCCGAGLLPEPDQRVYTCGGNGLGC 360
QY 361 GFAFCRECKEAYHEBECSAVFPAASGTTTQAYVDEBAAQAEWAASKETIKKTPCPR 420
DB 361 GFAFCRECKEAYHEBECSAVFPAASGTTTQAYVDEBAAQAEWAASKETIKKTPCPR 420
QY 421 CHVPEKNGGCMHMKCPQPCRLBWCNCGCEMNRVCMGDHMFV 465
DB 421 CHVPEKNGGCMHMKCPQPCRLBWCNCGCEMNRVCMGDHMFV 465
DB 393 CHVPEKNGGCMHMKCPQPCRLBWCNCGCEMNRVCMGDHMFV 437

RESULT 6
US-09-785-548-4
; Sequence 4, Application US/09785548
; Patent No. US20020155577A1
; GENERAL INFORMATION:
; APPLICANT: AVENTIS PHARMACEUTICALS, INC.
; TITLE OF INVENTION: COMPOSITIONS THAT CAN BE USED FOR REGULATING THE ACTIVITY OF PARKI
; FILE REFERENCE: ST00005
; CURRENT APPLICATION NUMBER: US/09/785,548
; CURRENT FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-785-548-4

Query Match          33.5%; Score 870; DB 9; Length 156;
Best Local Similarity 99.4%; Pred. No. 1.3e-68;
Matches 155; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```


QY 135 GSPAGRSIYNSFYVYCKGRCORVQPGKLRVOCSTCRQATLTLTGPGSCMDVLLPNNMSG 194
 DB 1 GSPAGRSIYNSFYVYCKGRCORVQPGKLRVOCSTCRQATLTLTGPGSCMDVLLPNNMSG 60
 QY 195 ECOSPHCPTSAEFPFKGCAHPTSDKETPVALHLIATNSRNITCTCTDVRSPLVAVQCN 254
 DB 61 ECOSPHCPTSAEFPFKGCAHPTSDKETPVALHLIATNSRNITCTCTDVRSPLVAVQCN 120
 QY 255 SRHVILCLOCFLYCYTRLNDROFVHDPOLGYSILPCV 290
 DB 121 SRHVILCLOCFLYCYTRLNDROFVHDPOLGYSILPCV 156

RESULT 7

US-10-239-249-2
 ; Sequence 2, Application US/10239249
 ; Publication No. US20030177507A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HONER, MARTUS
 ; APPLICANT: LINK, WOLFANG
 ; APPLICANT: BAUMEISTER, RALF
 ; TITLE OF INVENTION: NEMATODES AS MODEL ORGANISMS FOR INVESTIGATING
 ; TITLE OF INVENTION: NEURODEGENERATIVE DISEASES AND, IN PARTICULAR, PARKINSON'S
 ; TITLE OF INVENTION: DISEASE, USES AND METHODS FOR FINDING SUBSTANCES AND
 ; TITLE OF INVENTION: GENES WHICH CAN BE USED IN TREATING SUCH DISEASES, AND
 ; FILE REFERENCE: 02481.1804-00000
 ; CURRENT APPLICATION NUMBER: US/10/239,249
 ; CURRENT FILING DATE: 2002-09-20
 ; PRIOR APPLICATION NUMBER: PCT/EP01/03214
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 100 14 109.9
 ; PRIOR FILING DATE: 2000-03-22
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 386
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 ; US-10-239-249-2

Query Match
 Best Local Similarity 21.0%; Score 544.5; DB 14; Length 386;
 Matches 129; Conservative 59; Mismatches 166; Indels 93; Gaps 13;
 QY 22 SIFOLKEVVAARQGVADOLRVIFAGKELRNMTVONCULDQSIYHIVQPRKRGQENK 81
 DB 29 NIEDLTQDVEKLTETIPDELEVFCGKLSKSTIMEDSLTPAYQIMLRPFENSHNENG 88
 QY 82 ATGDDPRNNAAGGERPQSLTRVDLSSSVLPEDSVGLAVIILHTDSRKDPPAGSPAGRS 141
 DB 89 AT-----TAKITTS-----S 99
 QY 142 IYNSFYVYCKGRCORVQPGKLRVOCSTCRQATLTLTGPGSCMDVLLPNNMSGECQSPHC 201
 DB 100 ILGSPFYVYCKGN-CDDVYRGKLRVYCOCKSGSTVLVASEPQNNSDVLKSKRLPAVCEECCT 158
 QY 202 PGTSAEFPFKGCAHPTSDKETPVALHLIATNSRNITCTCTDVRSPLVAVFQCN SRHVILC 261
 DB 159 PGTFAPFKKCC-----LACNDPAALTHYRGWQMECCVCKQKQKVFIDLGN--HTTCQ 212
 QY 262 DCHLVCYTRLNDROFVHDPOLGYSILPC-VAGCPNSLKEHLHRIIGEEQYNNRYQOYGA 320
 DB 213 FCGRDVILSOLBRFGVQPHGFTIFCPYRQC-NRVQDVHNNHNGQTSSEYORKAT 271
 QY 321 EECV-LQMGVYCPRPGGAGLLEPP--DQRKVTCEGNGVLCGFAFCRECKEAYHGGC 377
 DB 272 ERLIYVDDKGVTCFVNSCGSGFWEFPYDDGRSQCP-----DGFSPCKKCC-----FERNC 322
 QY 378 SAVFEASGTTTQAVYVDERAABQARWEAASKETIKKTKRCPCGHYVPERNGSCMMKCP 437
 DB 323 VCOSHDLLTRT-----TIDATTRCPKCHVATERNNGCAHIHC- 360

QY 438 QPQRLIEWNCNCGENNRVCMGDHFPD 464
 DB 361 -TSCGMDWCFCKTEWKEBCQMDHFN 386

RESULT 8

US-10-313-203-10
 ; Sequence 10, Application US/10313203
 ; Publication No. US20040029134A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gu, Wei
 ; APPLICANT: Nikolaev, Anatoly
 ; TITLE OF INVENTION: P53-ASSOCIATED PARKIN-LIKE CYTOPLASMIC PROTEIN, AND RELATED COMPO
 ; FILE REFERENCE: 68106
 ; CURRENT APPLICATION NUMBER: US/10/313,203
 ; CURRENT FILING DATE: 2002-12-06
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 10
 ; LENGTH: 64
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-313-203-10

Query Match
 Best Local Similarity 14.4%; Score 373; DB 15; Length 64;
 Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 314 RYQOYGAEECYLQMGVLCPRPGCAGLLPEPDQKVTCEGNGVLCGFAFCRECKEAYH 373
 DB 1 RYQOYGAEECYLQMGVLCPRPGCAGLLPEPDQKVTCEGNGVLCGFAFCRECKEAYH 60
 QY 374 EGEC 377
 DB 61 EGEC 64

RESULT 9

US-10-313-203-17
 ; Sequence 17, Application US/10313203
 ; Publication No. US20040029134A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gu, Wei
 ; APPLICANT: Nikolaev, Anatoly
 ; TITLE OF INVENTION: P53-ASSOCIATED PARKIN-LIKE CYTOPLASMIC PROTEIN, AND RELATED COMPO
 ; FILE REFERENCE: 68106
 ; CURRENT APPLICATION NUMBER: US/10/313,203
 ; CURRENT FILING DATE: 2002-12-06
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 17
 ; LENGTH: 56
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-313-203-17

Query Match
 Best Local Similarity 12.4%; Score 323; DB 15; Length 56;
 Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 238 CITCTDVRSPLVAVFQCN SRHVILCLOCFLYCYTRLNDROFVHDPOLGYSILPCVAGC 293
 DB 1 CITCTDVRSPLVAVFQCN SRHVILCLOCFLYCYTRLNDROFVHDPOLGYSILPCVAGC 56

RESULT 10

US-09-864-761-36750
 ; Sequence 36750, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:34:50 / Search time 19.3575 Seconds
(without alignments)
2311.294 Million cell updates/sec

Title: US-10-622-817-9

Perfect score: 2596
Sequence: 1 MIVFRRFNSHGPFVPEVDS.....CMNCGCEWNRVCMGDMFEDV 465

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.79:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | ID | Description |
|------------|-------|---------------|--------|----------|--------------------|
| 1 | 490 | 18.9 | 357 | 2 T23460 | hypothetical prote |
| 2 | 220 | 8.5 | 543 | 2 H84724 | probable ARI-like |
| 3 | 219.5 | 8.5 | 1753 | 2 T00350 | hypothetical prote |
| 4 | 205 | 7.9 | 437 | 2 T16477 | hypothetical prote |
| 5 | 205 | 7.9 | 491 | 2 F87793 | protein C27A12.6 l |
| 6 | 198.5 | 7.6 | 644 | 2 T02366 | hypothetical prote |
| 7 | 196.5 | 7.6 | 514 | 2 G84724 | probable ARI-like |
| 8 | 195 | 7.5 | 320 | 2 T45909 | hypothetical prote |
| 9 | 193 | 7.4 | 542 | 2 H84725 | similar to Axiadhe |
| 10 | 190.5 | 7.3 | 436 | 2 H87793 | protein C27A12.8 l |
| 11 | 188 | 7.2 | 1048 | 2 T16553 | hypothetical prote |
| 12 | 186.5 | 7.2 | 324 | 2 T16983 | hypothetical prote |
| 13 | 182.5 | 7.0 | 551 | 2 S38086 | hypothetical prote |
| 14 | 180.5 | 6.9 | 594 | 2 T04783 | hypothetical prote |
| 15 | 178.5 | 6.9 | 497 | 2 G87793 | protein C27A12.7 l |
| 16 | 178.5 | 6.9 | 498 | 2 B86448 | hypothetical prote |
| 17 | 174.5 | 6.7 | 408 | 2 T47498 | hypothetical prote |
| 18 | 174 | 6.7 | 348 | 2 T47494 | hypothetical prote |
| 19 | 173 | 6.7 | 565 | 2 F84721 | probable RING zinc |
| 20 | 162.5 | 6.3 | 498 | 2 TC5983 | protein kinase C-1 |
| 21 | 159 | 6.1 | 518 | 2 E84536 | hypothetical prote |
| 22 | 158.5 | 6.1 | 22 | 2 T16481 | hypothetical prote |
| 23 | 157 | 6.0 | 488 | 2 T29562 | hypothetical prote |
| 24 | 155 | 6.0 | 532 | 2 T04748 | hypothetical prote |
| 25 | 152.5 | 5.9 | 816 | 2 T25555 | hypothetical prote |
| 26 | 149 | 5.7 | 464 | 2 S48329 | probable membrane |
| 27 | 147 | 5.7 | 1209 | 2 T52523 | ubiquitin / riboso |
| 28 | 143.5 | 5.5 | 155 | 2 JH0227 | ubiquitin / riboso |
| 29 | 142.5 | 5.5 | 155 | 2 JH0226 | ubiquitin / riboso |

| | | | | | |
|----|-------|-----|-----|----------|---------------------|
| 30 | 141.5 | 5.5 | 468 | 2 A84601 | Mutator-like trans |
| 31 | 136 | 5.2 | 222 | 2 T47487 | hypothetical prote |
| 32 | 134.5 | 5.2 | 229 | 1 UCUNI | ubiquitin precursor |
| 33 | 134 | 5.2 | 373 | 2 B84647 | hypothetical prote |
| 34 | 134 | 5.2 | 638 | 2 T02611 | hypothetical prote |
| 35 | 133 | 5.1 | 397 | 2 H84578 | probable RING zinc |
| 36 | 132.5 | 5.1 | 869 | 1 JC4858 | VLDL receptor prec |
| 37 | 132 | 5.1 | 80 | 2 T28305 | ORF MSV144 probabl |
| 38 | 132 | 5.1 | 150 | 2 T30390 | probable ubiquitin |
| 39 | 132 | 5.1 | 154 | 2 S55242 | polyubiquitin 2 - |
| 40 | 132 | 5.1 | 156 | 2 C86439 | protein T19E23.13 |
| 41 | 131 | 5.0 | 534 | 2 S34285 | polyubiquitin - to |
| 42 | 130.5 | 5.0 | 305 | 1 S29853 | polyubiquitin 4 - |
| 43 | 130.5 | 5.0 | 356 | 1 UCUTRC | polyubiquitin / r1 |
| 44 | 130 | 5.0 | 155 | 2 S40240 | ubiquitin/ribosome |
| 45 | 130 | 5.0 | 156 | 2 T52335 | ubiquitin extensio |

ALIGNMENTS

RESULT 1

T23460 hypothetical protein K083.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T23460
R:McMurray, A.
Submitted to the EMBL Data Library, November 1996
A:Reference number: T23460
A:Accession: T23460
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-357 <WILD>
A:Cross-References: UNIPROT:Q9XUS3; EMBL:281568; PIDN:CAB04599.1; GSPDB:GN00021; CESP:K01
A:Experimental source: clone K08E3
C:Genetics:
A:Gene: CESP:K08E3.7
A:Map position: 3
A:introns: 23/3; 72/2; 218/1; 286/2; 311/2

Query Match 18.9%; Score 490; DB 2; Length 357;
Best Local Similarity 26.7%; Pred. No. 1.8e-29;
Matches 119; Conservative 53; Mismatches 154; Indels 120; Gaps 12;

| | | | |
|----|-----|--|-----|
| Qy | 22 | STFOLKEVAVAKQGVPAQOLRYIFAGKELRNDWTQNCDDLQOQSVHYIVQREWRKQGEVN | 81 |
| Db | 29 | NIEDLTQVEKLTETPSDELVEVFCGKLSKSTIRDSLTPATOIMLRPFNSHNENG | 88 |
| Qy | 82 | ATGSDPPNNAAGCGRREPQSLTRVDLSSVLFPGDSVGLAVILHTPSRKDSPAAGSPAGRS | 141 |
| Db | 89 | AT-----TAKTTTS-----S | 99 |
| Qy | 142 | INSEFYVCKGPGQROPEKRLVQSCSTCRQATLTLTGSPSCWDVLIIPRMGEGCQSPHC | 201 |
| Db | 100 | ILGSFYVCKN-CDVVKRGLKRLVYQCKCSSTVVKSEFQNMVDLKSRIYAVCEBCT | 158 |
| Qy | 202 | PCTSAEFFPKCGAHPSTDETPVALHLIATNSRNITCTITVRSPLVLFQNSRNVICL | 261 |
| Db | 159 | PCLFAEFKFC---LACNDPAALTHVGNQMTECCDCDEKRVIPDLGCN--HITQ | 212 |
| Qy | 262 | DCFHLVCTRLNDRQFVHPDPLQVSLPCVAGSPNSLKEHLHFRILIGEEVRYOQYAE | 321 |
| Db | 213 | FCFRV-----RVQDVHFFHMGQSYSEYQKATE | 243 |
| Qy | 332 | ECV-LQMGVLCPRPGCGAGLPRP--DQRYTCGNGNLGCGAFRCBCKEAYHEGCS | 378 |
| Db | 244 | RLIAYDDKGVCPNNSCGGFFWEYDDGRGCP---DCFSPCRKC---FERNCV | 294 |
| Qy | 379 | AVFEASGTTQAYRVDEPAAEQARWEASKETIYKTTCPCPCHYVPEVNGGCMKCKQ | 438 |
| Db | 295 | COSEDDLTRT-----TIDATTRCPCHVATERNGCAGAHIC-- | 331 |

QY 439 POCRLWCNCGCENRVCMDHWD 464
DB 332 TSCGMDWCFCKCTEKKEECQMDHMFN 357

RESULT 2

probable ARI-like RING zinc finger protein (imported) - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: H847724
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentol, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffett, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.;
Nuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: H847724
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-543 <STO>
A:Cross-references: UNIPROT:Q9SKC3; GB:AE002093; NID:94887759; PIDN:ABD32295.1; GSPDB:GN
C:Genetics:
A:Gene: At2g31770
A:Map position: 2

Query Match 8.5%; Score 220; DB 2; Length 543;
Best Local Similarity 25.9%; Pred. No. 5.9e-09;
Matches 60; Conservative 33; Mismatches 83; Indels 56; Gaps 11;

QY 235 NITCITCTDVRSPVAFQCNHRVCLDCHLYCTRLND-----RQFVHPQGLSLPC 289
DB 125 NIQCGICFESTYREIARVSCGPHCYCTCMAGYITTKIEDPGCLRYKCEPS-----C 178
QY 290 VAGCENSLIKELHFRILIGSEQYNNY--QOYGAECVLOMGVLCPRPGGAGL---P 344
DB 179 SAAVKGMIEDVETKVV--NEKYSRYILRSYVEDGKKIKM-----CSPGCGYAVERGSS 232
QY 345 PDQKVTCEGNGAGCGAFCECKEAVHEG--ESAVFEASGTTQAYRDERAAR 403
DB 233 SSSYVSL-----CSTRFCMNCSEDAHSFVDDTV-----SKM 266
QY 404 -----EAAKERTIKTTKPCPRCHVEKNGGCMHKCPQPCRLEMCNCG 449
DB 267 IFKNQDESENKMWLANSKPCPECKRPIEKNDGCMHTGAP--CGHEFCMTC 317

RESULT 3

hypothetical protein KIAA0708 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C:Accession: T00350
R:Ichikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A:Reference number: Z14142; MUID:98403880; PMID:9734811
A:Accession: T00350
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1753 <ISH>
A:Cross-references: UNIPROT:Q8IWT3; EMBL:AB014608; NID:93327229; PIDN:BA411683.1; PID:93
A:Experimental source: brain
C:Genetics:
A>Note: KIAA0708

Query Match

Best Local Similarity 8.5%; Score 219.5; DB 2; Length 1753;
Matches 125; Conservative 57; Mismatches 181; Indels 221; Gaps 32;

QY 1 MIVPFRSSSGFPEV---DSST-----IFOLK 27
DB 1012 MWLTKKQTEVSVEITLLKSDSLSPDLLQAVLPLTSGNGPLTLHGQDFRGGVLRLLH 1071

QY 28 EVAKROG-----VPADLRVLPFAKETLNDMTV-----QNCDDDOOSI 66
DB 1072 EPGPORSGEALMLIPQAYLVNVEKDEGRITLQKRNLSCLVRLIKAHGKGLHTD--L 1129

QY 67 VHIYQPRKGGQEMNATGDDPRNNAAGCEREPOSITRVDSVLPFGDSVGLAVILHTD 126
DB 1130 VCVLTEMQKGPMPPTLG--HTVAGV-----ACTSTDVLSCTL--HLGGQVY---- 1175
QY 127 SRKDSPP-----AGSPAKRSIYNSFYVCKPGCRVOPGKRLRYCCSCRAATLTLYGP 180
DB 1176 KRSDRPPQILMYAAPFPWPCGQADVPCGSGSTSPSPBAV-----ATLASQLP 1228
QY 181 SCWDDVILNBRMSGEGSPHCSTAEFPFK--CGAHPSTDKETPVALLIATNS---- 233
DB 1229 -----AGRTMSFQ-----EVEGIMQYVRQVETLNLBPDAVQHLAHSHNGAE 1272

QY 234 -----RNITCITCTDVRSPV-----LVQCNHRV 259
DB 1273 QLLQSYSDPEPDLILAAGLVHQAQAVVRPDHCPVCV--SPUGCDDDLPSLC--CMHYC 1328
QY 260 CLDCHLYCTRLNDRQFVHPDQGLSLPC--VAGCP-----NSLIKELHFRILIGSEQYNR 314
DB 1329 CASKNNEYLTTRI-----EONLVINCTCPICADCPAQPTGAFIRAL--VSPVYISK 1377
QY 315 YQQ-----YGAEECVLOMGVLCPRP--GCCAGLLPEPDQKVTCEGNGILG----CGFA 363
DB 1378 YEKALLRGY--VESC--SNLTWCTNPGCD-----RILCRQGLCGYTCSCKGWA 1423
QY 364 FCRECK--EAMHEECGAVPEASGTTQAYRDERAAR 409
DB 1424 SCFNCSPFEAHYVASC-----GHMSQWDDOGGYDGKSVAPQSKH 1463

QY 410 TIKTTKPCPRCHVEKNGGCMHKCPQPCRLEMCNCGEW 453
DB 1464 LAKLISKRCSCQAPFKHKEGLHMTTC--AKCNHGCWCLXSW 1505

RESULT 4

hypothetical protein F56D2.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T16477
R:Du, Z.
submitted to the EMBL Data Library, August 1994
A:Description: The sequence of C. elegans cosmid F56D2.
A:Reference number: Z18519
A:Accession: T16477
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-437 <DUZ>
A:Cross-references: UNIPROT:Q20871; EMBL:U13644; NID:9532100; PID:91945502; PIDN:AA85268;
C:Genetics:
A:Gene: CESP.F56D2.5
A:Experimental source: strain Bristol N2; clone F56D2
A:Map position: 3
A:Introns: 41/3; 134/3; 185/3; 223/3; 254/2; 287/2; 414/3

Query Match

Best Local Similarity 7.9%; Score 205; DB 2; Length 437;
Matches 104; Conservative 66; Mismatches 169; Indels 176; Gaps 28;

QY 18 DSDTSIFOLKEVAVKRGVADQLRVLPFAKEL--RNDVTYQNCDDDOOSIVHIQRPWR 75
DB 3 DRDQIVEL-----EALSVLRKCKLANSDDMSDKAAET--QGIIYV----- 42
QY 76 KGGEMNATGDDPRNNAAGCEREPOSITRVDSVLPFGDSVGLAVILHTSRKXSPAG 135
DB 43 -----GPDNLVYDPTVITLGTSDGQFHLPLDILP-----PIRLKFLHPNDYPTVS 88
QY 136 SPAGRSIYNSFYVCKPGCRVOPGKRLRYCCSCRAATLTLYGSPCWDVILNBRMSGE 195

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 23, 2005, 13:33:04 ; Search time 88.371 Seconds
(without alignments)
2694.512 Million cell updates/sec

Title: US-10-622-817-9
Perfect score: 2596
Sequence: 1 MIVFVRFNSSHGPFVVDSD.....CWNCGCEMRVCGMDHMFV 465

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues
Total number of hits satisfying chosen parameters: 1612376

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_03:*
1: uniprot_prot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|--------------|--------------------|
| 1 | 2596 | 100.0 | 465 | 1 PRKN_HUMAN | O60260 homo sapien |
| 2 | 2218 | 85.4 | 465 | 1 PRKN_RAT | O91666 rattus norv |
| 3 | 2170.5 | 83.6 | 464 | 1 PRKN_MOUSE | O94966 mus musculu |
| 4 | 1115 | 43.0 | 203 | 2 OBN142 | O8n142 homo sapien |
| 5 | 1060 | 40.8 | 194 | 2 O95M07 | O95M07 bos taurus |
| 6 | 1058.5 | 40.8 | 482 | 2 O95T14 | O95T14 drosophila |
| 7 | 1058.5 | 40.8 | 482 | 2 O7KTX7 | O7KTX7 drosophila |
| 8 | 1049 | 40.4 | 514 | 2 O7O591 | O7O591 anopheles g |
| 9 | 1024 | 39.4 | 177 | 2 O8N141 | O8n141 homo sapien |
| 10 | 957 | 36.9 | 177 | 2 O8K5C2 | O8K5C2 rattus norv |
| 11 | 544.5 | 21.0 | 386 | 2 O9XUS3 | O9XUS3 caenorhabd |
| 12 | 543.5 | 20.9 | 162 | 2 O86LE7 | O86LE7 drosophila |
| 13 | 485 | 18.7 | 111 | 2 O8VHY5 | O8VHY5 rattus norv |
| 14 | 420.5 | 16.2 | 120 | 2 O7TPE8 | O7TPE8 mus musculu |
| 15 | 382 | 14.7 | 78 | 2 O8K1N5 | O8K1N5 mus musculu |
| 16 | 284 | 10.9 | 50 | 2 O6Q216 | O6Q216 homo sapien |
| 17 | 256 | 9.9 | 520 | 2 O6T486 | O6T486 dictyostell |
| 18 | 250.5 | 9.6 | 562 | 2 O84R80 | O84R80 arabidopsis |
| 19 | 249.5 | 9.6 | 562 | 2 O68R65 | O68R65 xenopus lae |
| 20 | 246 | 9.5 | 525 | 2 O653S8 | O653S8 oryza sativ |
| 21 | 240.5 | 9.3 | 503 | 1 AR11_DROME | O94581 drosophila |
| 22 | 240 | 9.2 | 492 | 2 O7T355 | O7T355 brachydanio |
| 23 | 240 | 9.2 | 511 | 2 O9VY94 | O9VY94 anopheles g |
| 24 | 238.5 | 9.1 | 527 | 2 O7OGU7 | O7OGU7 anopheles g |
| 25 | 236.5 | 9.1 | 445 | 2 O8CFJ4 | O8CFJ4 mus musculu |
| 26 | 236.5 | 9.1 | 533 | 2 O6NM85 | O6NM85 brachydanio |
| 27 | 236.5 | 9.1 | 555 | 1 AR11_MOUSE | O921K5 mus musculu |
| 28 | 236.5 | 9.1 | 557 | 1 AR11_HUMAN | O9Y4X5 homo sapien |
| 29 | 234.5 | 9.0 | 527 | 2 O6PFJ9 | O6PFJ9 brachydanio |
| 30 | 232.5 | 8.9 | 491 | 1 AR12_MOUSE | O921K6 mus musculu |
| 31 | 231 | 8.9 | 491 | 2 O6GL03 | O6GL03 xenopus tro |

| | | | | | |
|----|-------|-----|------|--------------|--------------------|
| 32 | 231 | 8.9 | 493 | 1 AR12_HUMAN | O95376 homo sapien |
| 33 | 231 | 8.9 | 607 | 2 O621V0 | O621V0 oryza sativ |
| 34 | 227.5 | 8.8 | 589 | 2 O7XRI6 | O7XRI6 oryza sativ |
| 35 | 225.5 | 8.7 | 501 | 2 O6CSC0 | O6CSC0 kluyveromyc |
| 36 | 225.5 | 8.7 | 2489 | 2 O68CP2 | O68CP2 homo sapien |
| 37 | 224 | 8.6 | 551 | 2 O84R11 | O84R11 arabidopsis |
| 38 | 224 | 8.6 | 552 | 2 O8L829 | O8L829 arabidopsis |
| 39 | 222 | 8.6 | 474 | 1 RN14_HUMAN | O9UB88 homo sapien |
| 40 | 222 | 8.6 | 474 | 2 O61BVO | O61BVO homo sapien |
| 41 | 222 | 8.6 | 492 | 2 O6RZU9 | O6RZU9 musa acumin |
| 42 | 220 | 8.5 | 40 | 2 O6S8G7 | O6S8G7 homo sapien |
| 43 | 220 | 8.5 | 543 | 2 O9SKC3 | O9SKC3 arabidopsis |
| 44 | 219.5 | 8.5 | 2517 | 1 PARC_HUMAN | O81WC3 homo sapien |
| 45 | 219.5 | 8.5 | 2517 | 2 O68D92 | O68D92 homo sapien |

ALIGNMENTS

RESULT 1
PRKN_HUMAN STANDARD: PRT: 465 AA.
ID PRKN_HUMAN
AC O60260; OBN143; OBN144; O9M07;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE parkin (EC 6.3.2.-) (Ubiquitin E3 ligase PRKN) (Parkinson juvenile disease protein 2) (Parkinson disease protein 2).
GN Name=PRKN2; Synonyms=PRKN;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND INVOLVEMENT IN JUVENILE PARKINSON'S DISEASE.
RC TISSUE=Fetal brain, and Skeletal muscle;
RX MEDLINE=98219084; PubMed=9560156; DOI=10.1038/33416;
RA Kitada T., Asakawa S., Hattori N., Matsunine H., Yamamura Y., Minoshima S., Yokochi M., Mizuno Y., Shimizu N.
RT "Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism." ;
RL Nature 392:605-608(1998).
RN
RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
RA D'Agata V., Scapagnini G., Cavallaro S.;
RT "Functional and molecular diversity of parkin." ;
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A. (ISOFORM 5).
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Shetty S.S., Loughran N.A., Peters G.J., Abmayr R.D., Mullen S.J., Bock S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richardson S., Morley K.C., Hale S.S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Bottinger A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield V.S.N., Krzywinski M.I., Skalski U., Smalls D.E., Scherch A., Schein J.E., Jones S.J.M., Maita M.A., full-length human and mouse cDNA sequences." ;
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN

RP FUNCTION IN UBIQUITINATION.
 RX PubMed=10973942; DOI=10.1074/jbc.C000447200;
 RA Imai Y., Soda M., Takahashi R.;
 RT "Parkin suppresses unfolded protein stress-induced cell death through
 RT its E3 ubiquitin-protein ligase activity.";
 RL N. J. Biol. Chem. 275:35661-35664(2000).
 RN [5]
 RP FUNCTION, AND CHARACTERIZATION OF VARIANTS PD PRO-42 AND ARG-240.
 RX PubMed=10888878; DOI=10.1038/77060;
 RA Shimura H., Hattori N., Kubo S.-I., Mizuno Y., Asakawa S.,
 RA Minoshima S., Shimizu N., Iwai K., Chiba T., Tanaka K., Suzuki T.;
 RT "Familial Parkinson disease gene product, parkin, is a ubiquitin-
 RT protein ligase.";
 RL Nat. Genet. 25:302-305(2000).
 RN [6]
 RP FUNCTION, AND INVOLVEMENT IN CANCER.
 RX PubMed=12719539; DOI=10.1073/pnas.0931262100;
 RA Cesari R., Martin E.S., Calin G.A., Pentimalli F., Bichi R.,
 RA McAdams H., Trapasso F., Drusco A., Shimizu M., Masciullo V.,
 RA D'Andrilli G., Scambia G., Picchio M.C., Alder H., Godwin A.K.,
 RT "Parkin, a gene implicated in autosomal recessive juvenile
 RT parkinsonism, is a candidate tumor suppressor gene on chromosome 6q25-
 RT q27.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:5956-5961(2003).
 RN [7]
 RP FUNCTION, INTERACTIONS WITH FBXW7 AND CUL1, AND UBIQUITINATION OF
 RX CYCLIN E.
 RA PubMed=12628165;
 RA Staropoli J.F., McDermott C., Martinat C., Schulman B., Demireva E.,
 RA Abellovich A.;
 RT "Parkin is a component of an SCF-like ubiquitin ligase complex and
 RT protects postmitotic neurons from kainate excitotoxicity.";
 RL Neuron 37:735-749(2003).
 RN [8]
 RP FUNCTION, UBIQUITINATION, AND S-NITROSYLATION.
 RX PubMed=15105460; DOI=10.1126/science.1093891;
 RA Chung K.K., Thomas B., Li X., Pletnikova O., Troncoso J.C., Marsh L.,
 RA Dawson V.L., Dawson T.M.;
 RT "S-nitrosylation of parkin regulates ubiquitination and compromises
 RT parkin's protective function.";
 RL Science 304:1328-1331(2004).
 RN [9]
 RP SUBCELLULAR LOCATION.
 RX PubMed=10319893;
 RA Shimura H., Hattori N., Kubo S.-I., Yoshikawa M., Kitada T.,
 RA Matsumine H., Asakawa S., Minoshima S., Yamamura Y., Shimizu N.,
 RA Mizuno Y.;
 RT "Immunohistochemical and subcellular localization of Parkin protein:
 RT absence of protein in autosomal recessive juvenile parkinsonism
 RT patients.";
 RL Ann. Neurol. 45:668-672(1999).
 RN [10]
 RP INTERACTIONS WITH UBE2L6 AND SEPT5, AND UBIQUITINATION OF SEPT5.
 RX PubMed=11078524; DOI=10.1073/pnas.240347997;
 RA Zhang Y., Gao J., Chung K.K., Huang H., Dawson V.L., Dawson T.M.;
 RT "Parkin functions as an E2-dependent ubiquitin-protein ligase and
 RT promotes the degradation of the synaptic vesicle-associated protein,
 RT CDCrel-1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:13354-13359(2000).
 RN [11]
 RP UBIQUITINATION OF GPR37.
 RX MEDLINE=21332586; PubMed=11439185; DOI=10.1016/S0092-8674(01)00407-X;
 RA Imai Y., Soda M., Inoue H., Hattori N., Mizuno Y., Takahashi R.;
 RT "An unfolded putative transmembrane polypeptide, which can lead to
 RT endoplasmic reticulum stress, is a substrate of Parkin.";
 RL Cell 105:851-902(2001).
 RN [12]
 RP INTERACTION, UBIQUITINATION OF SNCAIP, CHARACTERIZATION OF VARIANTS
 RP ARG-240, CYG-256, TRP-275 AND ASN-415, AND MUTAGENESIS OF
 RP CYG-337, CYG-421 AND CYG-431.
 RX PubMed=11590439; DOI=10.1038/hm1001-1144;
 RA Chung K.K., Zhang Y., Lim K.L., Tanaka Y., Huang H., Gao J.,
 RA

RA Ross C.A., Dawson V.L., Dawson T.M.;
 RT "Parkin ubiquitinates the alpha-synuclein-interacting protein,
 RT synphilin-1: implications for Lewy-body formation in Parkinson
 RT disease.";
 RL Nat. Med. 7:1144-1150(2001).
 RN [13]
 RP INTERACTIONS WITH STUB1 AND HSP70, AND UBIQUITINATION OF STUB1.
 RX PubMed=12150907;
 RA Imai Y., Soda M., Harakeyama S., Akagi T., Hashikawa T.,
 RA Nakayama K.-I., Takahashi R.;
 RT "CHIP is associated with Parkin, a gene responsible for familial
 RT Parkinson's disease, and enhances its ubiquitin ligase activity.";
 RL Mol. Cell 10:55-67(2002).
 RN [14]
 RP INTERACTION WITH SYT11, CHARACTERIZATION OF VARIANT PD GLY-289, AND
 RP MUTAGENESIS OF CYG-418
 RX PubMed=12925569; DOI=10.1093/hmg/ddg269;
 RA Huynh D.P., Soares D.R., Nguyen D., Pulst S.M.;
 RT "The autosomal recessive juvenile Parkinson disease gene product,
 RT parkin, interacts with and ubiquitinates synaptotagmin XI.";
 RL Hum. Mol. Genet. 12:2587-2597(2003).
 RN [15]
 RP UBIQUITINATION OF AN O-LINKED GLYCOSYLATED ISOFORM OF SNCAIP,
 RP SUBCELLULAR LOCATION, AND CHARACTERIZATION OF VARIANTS PD PRO-42 AND
 RP ARG-240.
 RX PubMed=11431533; DOI=10.1126/science.1060627;
 RA Shimura H., Schlossmacher M.G., Hattori N., Froesch M.P.,
 RA Trocenenbacher A., Schneider R., Mizuno Y., Kosik K.S., Selkoe D.J.;
 RT "Ubiquitination of a new form of alpha-synuclein by parkin from human
 RT brain: implications for Parkinson's disease.";
 RL Science 293:263-269(2001).
 RN [16]
 RP INVOLVEMENT IN CANCER.
 RX PubMed=14614460; DOI=10.1038/sj.onc.1207072;
 RA Denison S.R., Wang F., Becker N.A., Schuele B., Kock N.,
 RA Phillips L.A., Klein C., Smith D.I.;
 RT "Alterations in the common fragile site gene Parkin in ovarian and
 RT other cancers.";
 RL Oncogene 22:8370-8378(2003).
 RN [17]
 RP REVIEW.
 RX PubMed=15229644; DOI=10.1038/sj.embor.7400188;
 RA Kahle P.J., Haass C.;
 RT "How does parkin ligase ubiquitin to Parkinson's disease?";
 RL EMBO Rep. 5:681-685(2004).
 RN [18]
 RP STRUCTURE BY NMR OF 1-76, AND INTERACTION WITH PSMD4.
 RX PubMed=12634850; DOI=10.1038/sj.embor.embor764;
 RA Sakata E., Yamaguchi Y., Kurimoto E., Kikuchi J., Yokoyama S.,
 RA Yamada S., Kawahara H., Yokosawa H., Hattori N., Mizuno Y., Tanaka K.,
 RA Kato K.;
 RT "Parkin binds the Rpn10 subunit of 26S proteasomes through its
 RT ubiquitin-like domain.";
 RL EMBO Rep. 4:301-306(2003).
 RN [19]
 RP REVIEW ON VARIANTS.
 RX PubMed=14976155; DOI=10.1093/hmg/ddh089;
 RA Mata I.F., Lockhart P.J., Farrer M.J.;
 RT "Parkin genetics: one model for Parkinson's disease.";
 RL Hum. Mol. Genet. 13:R127-R133(2004).
 RN [20]
 RP VARIANT ARG-JP ARG-240.
 RX PubMed=9731209;
 RA Hattori N., Matsumine H., Asakawa S., Kitada T., Yoshino H.,
 RA Elibol B., Brookes A.J., Yamamura Y., Kobayashi T., Wang M.,
 RA Yoritaka A., Minoshima S., Shimizu N., Mizuno Y.;
 RT "Point mutations (Thr240Arg and Glu31Stop) in the Parkin gene.";
 RL Biochem. Biophys. Res. Commun. 249:754-758(1998).
 RN [21]
 RP ERRATUM.
 RA Hattori N., Matsumine H., Asakawa S., Kitada T., Yoshino H.,
 RA Elibol B., Brookes A.J., Yamamura Y., Kobayashi T., Wang M.,
 RA Yoritaka A., Minoshima S., Shimizu N., Mizuno Y.;

RL Biochem. Biophys. Res. Commun. 251:666-666(1998).
RN [122]
Query Match 100.0%; Score 2596; DB 1; Length 465;
Best Local Similarity 100.0%; Pred. No. 1,3e-195;
Matches 465; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MIVFRPNSSHGFPVEVDSDTISIPLKKEVAVRQGVADQLAVIPAGKELRNDMTVQND 60
DB 1 MIVFRPNSSHGFPVEVDSDTISIPLKKEVAVRQGVADQLAVIPAGKELRNDMTVQND 60
QY 61 LDQGGIVHIVQRPWRKGGEMATGDDPRNAAAGCEREPOSJTRVDSLSSVLPDVGIA 120
DB 61 LDQGGIVHIVQRPWRKGGEMATGDDPRNAAAGCEREPOSJTRVDSLSSVLPDVGIA 120
QY 121 VILHTDSKRDSPAPASGRSITNSFYVCKPCCQVQPGKLRVQSTCRQATLTLTQGP 180
DB 121 VILHTDSKRDSPAPASGRSITNSFYVCKPCCQVQPGKLRVQSTCRQATLTLTQGP 180
QY 181 SCMDVLLPNRMSGSCSPHCPGTSAEPFKCAHPTSKETPVALLHLATNSRNTCTT 240
DB 181 SCMDVLLPNRMSGSCSPHCPGTSAEPFKCAHPTSKETPVALLHLATNSRNTCTT 240
QY 241 CTDVRSPLYVPCQNSRHVICTLDCPHLYCVTRLNDQFVHDPLQGYSLPCVAGCPNLSIKE 300
DB 241 CTDVRSPLYVPCQNSRHVICTLDCPHLYCVTRLNDQFVHDPLQGYSLPCVAGCPNLSIKE 300
QY 301 LHHFPIILBEOYNRVOQGAEECVLQMGVLCPRPGCGAGLLPEPDQRYTEGGNGLGC 360
DB 301 LHHFPIILBEOYNRVOQGAEECVLQMGVLCPRPGCGAGLLPEPDQRYTEGGNGLGC 360
QY 361 GPAPRECKEAVHEBEGSCAVFEASGTTTQAYRVDEBAARQARWEAASKETIKTKPCPR 420
DB 361 GPAPRECKEAVHEBEGSCAVFEASGTTTQAYRVDEBAARQARWEAASKETIKTKPCPR 420
QY 421 CHVPEKNGGCMHMKCPOPOCRLENCWNGCEMNRVCMGDHMFV 465
DB 421 CHVPEKNGGCMHMKCPOPOCRLENCWNGCEMNRVCMGDHMFV 465
RESULT 2
PRKN RAT STANDARD; PRT; 465 AA.
ID _PRKN RAT STRAND; PRT; 465 AA.
AC 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Parkin (BC 6.3.2.-) (Ubiquitin E3 ligase PRKN).
GN Name=Park2; Synonyms=Prkn;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4; 5 AND 6).
RC STRAIN=Sprague-Dawley;
RX MEDLINE=20153712; PubMed=10686358; DOI=10.1016/S0169-328X(99)00286-7;
RA D'Agata V., Zhao W., Cavallaro S.,
RT "Cloning and distribution of the rat parkin mRNA."
RL Brain Res. Mol. Brain Res. 75:345-349(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE SPECIFICITY.
RC TISUE=Hypothalamus;
RX MEDLINE=20199898; PubMed=10737637;
RA Gu W.-D., Abbas N., Lagunes M.Z., Parent A., Pradler L., Bohne G.A.,
RT Agid Y., Hirsch E.C., Ragsman-Vozari R., Brice A.,
RL "Cloning of rat parkin cDNA and distribution of parkin in rat brain."
RN J. Neurochem. 74:1773-1776(2000).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=Sprague-Dawley;
RA Hattori N., Wang M., Mizuno Y.,
RT "The expression of parkin mRNA in developing, adult and ageing rat

RT CNS."?
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RA Soda M., Imai Y., Takahashi R.,
RT "Molecular cloning of rat parkin gene."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Functions within a multiprotein E3 ubiquitin ligase
complex, catalyzing the covalent attachment of ubiquitin moieties
onto substrate proteins. These substrates include SYTL1, CME1,
GPR37, STUB1, a 22 kDa O-linked glycosylated isoform of SNCAIP and
SEPT5. May play a more general role in the ubiquitin proteasomal
pathway by participating in the removal and/or deoxidification of
abnormally folded or damaged protein. Loss of this ubiquitin
ligase activity appears to be the mechanism underlying
pathogenesis of AR-OP. May protect neurons against alpha synuclein
toxicity, proteasomal dysfunction, GPR37 accumulation, and
kainate-induced excitotoxicity. May play a role in controlling
neurotransmitter trafficking at the presynaptic terminal and in
calcium-dependent exocytosis. Regulates cyclin B during neuronal
apoptosis. May represent a tumor suppressor gene (By similarity).
CC -1- PATHWAY: Ubiquitin conjugation; third step.
CC -1- SUBUNIT: Forms an E3 ubiquitin ligase complex with UBE2L3 or
UBE2L6. Part of a SCF-like complex, consisting of PARK2, CUL1 and
FBXW7. Interacts with SNCAIP. Binds to the C2A and C2B domains of
SYTL1. Interacts and regulates the turnover of SEPT5. Part of a
complex, including STUB1, HSP70, and GPR37. The amount of STUB1 in
the complex increases during ER stress. STUB1 promotes the
disassociation of HSP70 from PARK2 and GPR37, thus facilitating
PARK2-mediated GPR37 ubiquitination. HSP70 transiently associates
with unfolded GPR37 and inhibits the E3 activity of PARK2,
whereas, STUB1 enhances the E3 activity of PARK2 through promotion
of disassociation of HSP70 from PARK2-GPR37 complexes. Interacts
with PSMD4 (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Expressed in the endoplasmic
reticulum, dendrites, some presynaptic terminals and in
postsynaptic densities (By similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=6;
CC Name=1;
CC IsoId=Q9JK66-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9JK66-2; Sequence=VSP_011722, VSP_011723;
CC Name=3;
CC IsoId=Q9JK66-3; Sequence=VSP_011717;
CC Name=4;
CC IsoId=Q9JK66-4; Sequence=VSP_011718;
CC Name=5;
CC IsoId=Q9JK66-5; Sequence=VSP_011719;
CC Name=6;
CC IsoId=Q9JK66-6; Sequence=VSP_011717, VSP_011720, VSP_011721;
CC -1- TISSUE SPECIFICITY: Largely confined to neuronal elements,
including fibers and neuropil. Highly expressed at the forebrain
level, in pyramidal cells of layer V, in various cortical regions
and cerebellum. Expressed in the nucleus of diagonal band of
Broca, nucleus basalis, bed nucleus of the stria terminalis, and
olfactory tubercle. Moderate expression is seen in most neurons of
the subthalamic nucleus, heart, skeletal muscle and testis.
CC Moderate expression was found in frontal cortex, parietal cortex,
cerebellum, heart, skeletal muscle and testis.
CC -1- DOMAIN: The ubiquitin-like domain binds the PSMD4 subunit of 26S
proteasome (By similarity).
CC -1- PTM: Auto-ubiquitinates in an E2-dependent manner leading to its
own degradation (By similarity).
CC -1- PTM: S-nitrosylated (By similarity).
CC -1- SIMILARITY: Contains 2 IBR-type zinc fingers.
CC -1- SIMILARITY: Contains 1 ubiquitin-like domain.
CC -1- CAUTION: Has been said to contain 2 RING fingers, but these are
not found by any domain detection methods.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

family1 Parkinson's disease";

5. Biomat. NMR 25:153-156(2003).

-1- FUNCTION: Functions within a multiprotein E3 ubiquitin ligase complex, catalyzing the covalent attachment of ubiquitin moieties onto substrate proteins. These substrates include SYT11, CNER1, GPR37, STUB1, a 22 kDa O-linked glycosylated isoform of SNCAIP and SEPT5. May play a more general role in the ubiquitin proteasomal pathway by participating in the removal and/or detoxification of abnormally folded or damaged protein. Loss of this ubiquitin ligase activity appears to be the mechanism underlying pathogenesis of AR-JP. May protect neurons against alpha synuclein toxicity, proteasomal dysfunction, GPR37 accumulation, and kainate-induced excitotoxicity. May play a role in controlling neurotransmitter trafficking at the presynaptic terminal and in calcium-dependent exocytosis. Regulates cyclin E during neuronal apoptosis. May represent a tumor suppressor gene (By similarity).

-1- PATHWAY: Ubiquitin conjugation; third step.

-1- SUBUNIT: Forms an E3 ubiquitin ligase complex with UBE2L3 or UBE2L6. Part of a SCF-like complex, consisting of PARK2, CUL1 and FBXW7. Interacts with SNCAIP. Binds to the C2A and C2B domains of SYT11. Interacts and regulates the turnover of SEPT5. Part of a complex, including STUB1, HSP70, and GPR37. The amount of STUB1 in the complex increases during ER stress. STUB1 promotes the dissociation of HSP70 from PARK2 and GPR37, thus facilitating PARK2-mediated GPR37 ubiquitination. HSP70 transiently associates with unfolded GPR37 and inhibits the E3 activity of PARK2, whereas, STUB1 enhances the E3 activity of PARK2 through promotion of dissociation of HSP70 from PARK2-GPR37 complexes. Interacts with PSMD4 (By similarity).

-1- SUBCELLULAR LOCATION: Cytoplasmic. Expressed in the endoplasmic reticulum, dendrites, some presynaptic terminals and in postsynaptic densities.

-1- ALTERNATIVE PRODUCTS:

Event=Alternative splicing; Named Isoforms=3;

Name=1;

IsoId=G9WVS6-1; Sequence=Displayed;

Name=2;

IsoId=G9WVS6-2; Sequence=VSP_011714, VSP_011715;

Name=3;

IsoId=G9WVS6-3; Sequence=VSP_011713, VSP_011716;

-1- TISSUE SPECIFICITY: Expressed in all subdivisions of the brain. Highly expressed in brainstem, cranial nerve, pons, cerebellar nuclei, indusium griseum, nuclei reticularis, strata optica and lacunosum moleculare of the hippocampal CA2 region. Low levels were found in the telencephalon and diencephalon. Expressed in heart, liver, skeletal muscle, kidney and testis.

-1- DEVELOPMENTAL STAGE: In late 10 dpc weakly expressed in postmitotic neurons in the mantle layer of the developing nervous system. Expression increased at 11-12 dpc. At 15-16 dpc, as more specialized neurons and nonneuronal cells are formed, expression is more tissue specific. Expression was highest in the neurites, moderate levels were observed in the migrating postmitotic neurons in the intermediate and neopallial layers. In the diencephalon and other CNS regions, while the weakest level of expression was observed in the cell bodies. In nonneuronal tissues, high levels of expression were found in the muscle walls of the intestine, the blood vessels and the dermis.

-1- DOMAIN: The ubiquitin-like domain binds the PSMD4 subunit of 26S proteasomes (By similarity).

-1- PTM: Auto-ubiquitinates in an E2-dependent manner leading to its own degradation (By similarity).

-1- PTM: S-nitrosylated.

-1- SIMILARITY: Contains 2 IBR-type zinc fingers.

-1- SIMILARITY: Contains 1 ubiquitin-like domain.

-1- CAUTION: Has been said to contain 2 RING fingers, but these are not found by any domain detection methods.

This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use. By non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>)

| | | |
|---------------------------|--|--|
| CC | | or send an email to license@ib-ib.ch . |
| DR | EMBL; AB019558; BAAB2404.1; | - |
| DR | EMBL; AF250293; AAGJ3890.1; | - |
| DR | EMBL; AF250294; AAGJ3891.1; | - |
| DR | EMBL; AF250295; AAGJ3892.1; | - |
| DR | PDB; 1MG8; NMR; A=1-76. | |
| DR | MGI; MGI:1355296; Park2. | |
| DR | InterPro; IPR003977; parkin. | |
| DR | InterPro; IPR006262; Ubiquitin. | |
| DR | InterPro; IPR002867; znf_C6HC. | |
| DR | pfam; PF01485; IBR; 1. | |
| DR | pfam; PF00240; ubiquitin; 1. | |
| DR | PRINTS; PR01458; PARKIN. | |
| DR | PRINTS; PR00348; UBIQUITIN. | |
| DR | SMART; SMO0647; IBR; 2. | |
| DR | SMART; SMO0213; UBO; 1. | |
| DR | PROSITE; PS00299; UBIQUITIN_1; FALSE_NEG. | |
| KW | 3D-structure; Alternative splicing; Ligase; Metal-binding; Repeat; | |
| KW | S-nitrosylation; Ub1 conjugation; Ub1 conjugation pathway; Zinc; | |
| KW | Zinc-finger. | |
| FT | DOMAIN | 1 76 Ubiquitin-like. |
| FT | DOMAIN | 204 238 StyI binding 1. |
| FT | DOMAIN | 257 293 StyII binding 2. |
| FT | ZN_FING | 313 377 IBR-type 1. |
| FT | ZN_FING | 400 456 IBR-type 2. |
| FT | VARSPLIC | 244 261 RSPVLVEQCNRRHYICLD -> SHLPSSGASVWTRPHLH (in isoform 3). |
| FT | VARSPLIC | /FTId=VSP_011713. |
| FT | VARSPLIC | 245 254 SPVLVFQCNH -> FPKNSKRTS (in isoform 2). |
| FT | VARSPLIC | 255 464 /FTId=VSP_011714. |
| FT | VARSPLIC | Missing (in isoform 2). |
| FT | VARSPLIC | /FTId=VSP_011715. |
| FT | VARSPLIC | 262 464 Missing (in isoform 3). |
| FT | CONFLICT | 137 137 /FTId=VSP_011716. |
| FT | SEQUENCE | 464 AA; 51617 MW; 5574A285A9AIB080 CRC64; |
| Query Match | | 83.6%; Score 2170.5; DB 1; Length 464; |
| Beat Local Similarity | | 83.4%; Pred. No. 3,6e-162; |
| Matches 388; Conservative | | 29; Mismatches 47; Indels 1; Gaps 1, |
| OY | 1 | MI VFVRFNSSHGFVEVDSDTSIFOLKEVNAKRQGVPAADQLRVIPAKGEKLRNDMTVQNCD 60 |
| Db | 1 | MI VFVRFNSSHGFVEVDSDTSILQLEKVNAKRQGVPAADQLRVIPAKGEKLRNDMTVQNCD 60 |
| OY | 61 | LDQGSIVHIIVRPMPRKGEKNNAATGGDDPRNAAAGCEREPOSLTVDVDLSSTVLRGDVSGLA 120 |
| Db | 61 | LEOQSIVHIIVRPMPRKSHETNASGDEDPQSISEBSIWESRSRLTVDVDLSHTLVPDVSGLA 120 |
| OY | 121 | VILHRDSKSDPPPGSGPAGRSIVNSFFYYCYCKGPCORVOVGKLRLVOCSTCRQATLTLTQGP 180 |
| Db | 121 | VILDIDSKDEBAAARGPV-KPTYSFFLYCGPGCHKVOPGKLRLVOCSTCRQATLTLTQGP 179 |
| OY | 181 | SCMDVLIIPNMSEGCOSPHCPGTSAEFFFFKCGAHPTSDKETPVALLHLIATNSNITCIT 240 |
| Db | 180 | SCMDVLIIPNMSEGCOSPDCGRIRAEFFFCGAHPISDKOTSVALNLITSNRSSIPCLA 239 |
| OY | 241 | CTDVRSPLYLVFQCNSRHVICTLDCHLYCVTRLNDROPYHDPLQGLSYLPCVAGCENSLIKE 300 |
| Db | 240 | CTDVRSPLYLVFQCNSRHVICTLDCHLYCVTRLNDROPYHDQGLSYLPCVAGCENSLIKE 299 |
| OY | 301 | LHHFRPIIGEEQYNVYOYGAEECYLQMGCVLCPRPGCAGLLPEPDQRKYTCBEANGLG 360 |
| Db | 300 | LHHFRPIIGEEQYTRYOYGAEECYLQMGCVLCPRPGCAGLLPEPDQRKYTCBEANGLG 359 |
| OY | 361 | GFAFCREKEAYHGEGCSAVFEASGTTTQAARVDERAAEQARWEASAASKTIKTTKTCPR 420 |
| Db | 360 | GFVFCROCKEAYHBGDDCDLSLEPGASTSQAAARVDRKAABAQARWEASAASKTIKTTKTCPR 419 |
| OY | 421 | CHVPERKNGSGMHNKCPOPCRLEMCWNNGCENNRVCMDHWFDV 465 |

Db 420 CNVPIEKNGGCMHKPCQPQPCXKLEMCWNCGCCENRACMBDHWDPV 464

RESULT 4

Q8N142 PRELIMINARY; PRT; 203 AA.

ID Q8N142
AC Q8N142;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Parkin isoform.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA D'Agata V., Scapagnini G., Cavallaro S.;
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF381284; AAM21459.1; -
DR GO; GO:0004842; Publigitin-protein ligase activity; IEA.
DR GO; GO:0006512; Publigitin cycle; IEA.
DR InterPro: IPR003977; Parkin.
DR InterPro: IPR002867; Znf_C6HC.
DR Pfam: PF01485; IER; 1.
DR PRINTS; PR01475; PARKIN.
DR SMART; SM00647; IBR; 1.
SQ SEQUENCE 203 AA; 22192 MW; F92E796F069DC818 CRC64;

Query Match 43.0%; Score 1115; DB 2; Length 203;
Best Local Similarity 100.0%; Pred. No. 1.2e-79;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 192 MSGECQSHPCGTSAEFPFKCGAHPSTSKPEFVALHLATNSRNITCITCTDVSPPVLF 251
Db 1 MSGECQSHPCGTSAEFPFKCGAHPSTSKPEFVALHLATNSRNITCITCTDVSPPVLF 60

QY 252 QCNRRHYVCLDPCFHLVCTRLNDROFVHDPOLGYSLPCVACCPNSLIXELHHFRITGEEQ 311
Db 61 QCNRRHYVCLDPCFHLVCTRLNDROFVHDPOLGYSLPCVACCPNSLIXELHHFRITGEEQ 120

QY 312 YNRVQYGAECVYOMGSGVLPDPKCGAGLLPEPDRKYVTEGGNGLGGGAFRCCKEA 371
Db 121 YNRVQYGAECVYOMGSGVLPDPKCGAGLLPEPDRKYVTEGGNGLGGGAFRCCKEA 180

QY 372 YHEGCSAVFEASGTTQ 389
Db 181 YHEGCSAVFEASGTTQ 198

RESULT 5

Q95M07 PRELIMINARY; PRT; 194 AA.

ID Q95M07
AC Q95M07;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Parkin (Fragment).
GN Name=Parkin gene;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Brain;
MEDLINE=98219084; PubMed=9560156; DOI=10.1038/33416;
Kitada T., Asakawa S., Hattori N., Matsumine H., Yamamura Y.,
Minoshima S., Yokochi M., Mizuno Y., Shimizu N.;
RT "Mutations in the parkin gene cause autosomal recessive juvenile
parkinsonism".
RL Nature 392:605-608 (1998).

| | | |
|--|---|---|
| RN | [12] | SEQUENCE FROM N.A. |
| RP | TISSUE=Brain. | |
| RC | MEDLINE=21869006; PubMed=11879804; DOI=10.1016/S0006-8993(02)02241-2; | |
| RA | Hase A., Yamada H., Arai K., Sunada Y., Shimizu T., Matsumura K.; | |
| RT | "Characterization of parkin in bovine peripheral nerve." | |
| RL | Brain Res. 930:143-149(2002). | |
| DR | EMBL; AB060701; BAB70670.1; - | |
| DR | GO; GO:0004842; Fubiquitin-protein ligase activity; IEA. | |
| DR | GO; GO:0006512; P ubiquitin cycle; IEA. | |
| DR | InterPro; IPR003977; parkin. | |
| FR | PRINTS; PRO1475; PARKIN. | |
| FT | NON_TER | 1 |
| FT | NON_TER | 1 |
| SQ | SEQUENCE | 194 AA; 21225 MW; 9B214F09FA74FAB1 CRC64; |
| | | |
| Query Match | 40.8%; | Score 1060; DB 2; Length 194; |
| Best Local Similarity | 99.5%; | Pred.No. 2.4e-75; |
| Matches 193; Conservative | 0; | Mismatches 1; Indels 0; Gaps 0 |
| | | |
| QY | 92 AGCEREPOSILTRVYLSSSVPGDSVGAVLIHTDSDRSPAGSPAGSTINSFYVCK | 151 |
| Db | 1 AGGGEREQSILTRVDLSSSVLPGBDSVGLAVLIHTDSDRSPAGSPAGSTINSFYVCK | 60 |
| QY | 152 GPGRVPGKLRVOCSTRQAATLTITGPSAMDVLIPNRSGEQSPHCPGTSAEFFEK | 211 |
| Db | 61 GPCCRVPGKLRVOCSTRQAATLTITGPSAMDVLIPNRSGEQSPHCPGTSAEFFEK | 120 |
| QY | 212 CGAHPTSKETPVALLIATNSRNITCTCDIVSPVVPFCNSRHVICLCDFHLVCYTR | 271 |
| Db | 121 CGAHPTSKETPVALLIATNSRNITCTCTDVSPVLPFCNSRHVICLCDFHLVCYTR | 180 |
| QY | 272 LNDRFVHDPLQGY 285 | |
| Db | 181 LNDRFVHDPLQGY 194 | |
| | | |
| RESULT 6 | | |
| Q95T14 | PRELIMINARY; | PRT; 482 AA. |
| AC Q95T14; | | |
| DT 01-DEC-2001 (TREMBLrel. 19, Created) | | |
| DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update) | | |
| DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update) | | |
| DE SDOI679P. | | |
| CN Name=Park; Synonyms=Cg10523; | | |
| OS Drosophila melanogaster (fruit fly). | | |
| OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; | | |
| OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; | | |
| OC Ephydroidea; Drosophilidae; Drosophila. | | |
| OX NCBI_TaxId=7227; | | |
| FN [1] | | |
| RE SEQUENCE FROM N.A. | | |
| RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J., | | |
| RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R., | | |
| RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J., | | |
| RA Nunco J., Paclob J., Paragas V., Park S., Phouanavong S., Wan K., | | |
| RA Yu C., Lewis S.E., Rubin G.M., Ceinker S.; | | |
| Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases. | | |
| RL EMBL; AY058754; AAL13983.1; - | | |
| DR HSSP; Q862M4; IAAR. | | |
| DR FLYBase; FBgn0041100; park. | | |
| DR GO; GO:0007005; Mitochondrion organization and biogenesis; IMP. | | |
| DR InterPro; IPR002052; N6 Mtase. | | |
| DR InterPro; IPR003977; parkin. | | |
| DR InterPro; IPR000626; Ubiquitin. | | |
| DR InterPro; IPR002867; Znf_C6HC. | | |
| DR Pfam; PF01485; IBR; 1. | | |
| DR Pfam; PF00240; ubiquitin; 1. | | |
| DR PRINTS; PRO1475; PARKIN. | | |
| DR PRINTS; PRO0348; UBQUITIN. | | |
| DR SMART; SM00647; IBR; 2. | | |
| DR SMART; SM00213; UBO; 1. | | |

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: February 25, 2005, 16:48:21 ; Search time 2103.69 Seconds
(without alignments)
10848.741 Million cell updates/sec

Title: US-10-622-817-10
Perfect score: 471
Sequence: 1 ggaagctccagcagctagatc.....accctcctgctgtcgtag 471

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:
1: gb_ba:*
2: gb_bcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_gy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 471 | 100.0 | 471 | 6 | AX224425 Sequence |
| 2 | 468 | 99.4 | 2960 | 6 | BD269536 Mutations |
| 3 | 466.4 | 99.0 | 904 | 6 | AF381283 Homo sapi |
| 4 | 466.4 | 99.0 | 2960 | 6 | AR492172 Sequence |
| 5 | 466.4 | 99.0 | 2960 | 6 | AX584224 Sequence |
| 6 | 466.4 | 99.0 | 2960 | 6 | AX584226 Sequence |
| 7 | 466.4 | 99.0 | 2960 | 6 | AB009973 Homo sapi |
| 8 | 464.8 | 98.7 | 1575 | 9 | BC022014 Homo sapi |
| 9 | 460.8 | 97.8 | 1157 | 9 | AF381282 Homo sapi |
| 10 | 454 | 96.4 | 583 | 4 | AB060701 Bos tauru |
| 11 | 353.4 | 75.0 | 1564 | 10 | AB039878 Rattus no |
| 12 | 350.2 | 74.4 | 904 | 10 | AF343575 Rattus no |
| 13 | 348.6 | 74.0 | 1466 | 10 | AF168004 Rattus no |
| 14 | 348.6 | 74.0 | 1576 | 10 | AF210434 Rattus no |
| 15 | 348.2 | 73.9 | 1157 | 10 | AF343574 Rattus no |
| 16 | 347 | 73.7 | 1398 | 10 | AF257234 Rattus no |
| 17 | 347 | 73.7 | 1470 | 10 | AF381277 Rattus no |
| 18 | 347 | 73.7 | 1554 | 10 | AF381280 Rattus no |
| 19 | 347 | 73.7 | 1578 | 10 | AF381278 Rattus no |

| | | | | | |
|----|-------|------|------|----|--------------------|
| 20 | 340.2 | 72.2 | 1644 | 10 | AB019558 Mus muscu |
| 21 | 340.2 | 72.2 | 3043 | 6 | AX089875 Sequence |
| 22 | 340.2 | 72.2 | 3043 | 6 | AX136840 Sequence |
| 23 | 340.2 | 72.2 | 3092 | 6 | AX089879 Sequence |
| 24 | 340.2 | 72.2 | 3092 | 6 | AX136844 Sequence |
| 25 | 340.2 | 72.2 | 3226 | 10 | AF250293 Mus muscu |
| 26 | 340.2 | 72.2 | 3253 | 6 | AX089876 Sequence |
| 27 | 340.2 | 72.2 | 3253 | 6 | AX089878 Sequence |
| 28 | 340.2 | 72.2 | 3253 | 6 | AX136841 Sequence |
| 29 | 340.2 | 72.2 | 3253 | 6 | AX136843 Sequence |
| 30 | 340.2 | 72.2 | 3254 | 6 | AX089877 Sequence |
| 31 | 340.2 | 72.2 | 3254 | 6 | AX136842 Sequence |
| 32 | 340.2 | 72.2 | 3255 | 6 | AX089863 Sequence |
| 33 | 340.2 | 72.2 | 3255 | 6 | AX089880 Sequence |
| 34 | 340.2 | 72.2 | 3255 | 6 | AX089881 Sequence |
| 35 | 340.2 | 72.2 | 3255 | 6 | AX089882 Sequence |
| 36 | 340.2 | 72.2 | 3255 | 6 | AX136828 Sequence |
| 37 | 340.2 | 72.2 | 3255 | 6 | AX136845 Sequence |
| 38 | 340.2 | 72.2 | 3255 | 6 | AX136846 Sequence |
| 39 | 340.2 | 72.2 | 3255 | 6 | AX136847 Sequence |
| 40 | 338.6 | 71.9 | 3014 | 6 | AX089869 Sequence |
| 41 | 338.6 | 71.9 | 3014 | 6 | AX136834 Sequence |
| 42 | 338.4 | 71.8 | 1158 | 9 | AF381284 Homo sapi |
| 43 | 288.4 | 61.2 | 2876 | 6 | AR492173 Sequence |
| 44 | 265.4 | 56.3 | 3136 | 6 | AX089872 Sequence |
| 45 | 265.4 | 56.3 | 3136 | 6 | AX136837 Sequence |

ALIGNMENTS

RESULT 1
LOCUS AX224425 471 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 3 from Patent WO0160857.
ACCESSION AX224425
VERSION AX224425.1 GI:15554665
KEYWORDS

SOURCE
ORGANISM Homo sapiens (human)

REFERENCE
1 Koultnikova,H., Brice,A., Fournier,A., Pradier,L., Prades,C.,
Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS
1 Arnould-Reguigne,I., Rosier-Montus,M.F. and Corti,O.
Compositions useful for regulating parkin gene activity
Patent: WO 0160857-A 3 23-AUG-2001;
Aventis Pharma S.A. (FR) ; INSTITUT NATIONAL DE LA SANTE ET DE LA
RECHERCHE MEDICALE (INSERM) (FR)

TITLE
JOURNAL
JOURNAL

FEATURES
Source
Location/Qualifiers
1..471
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
<1..471
/note="unamed protein product"

CDS
/codon_start=1
/protein_id="CAC69572.1"
/db_xref="GI:15554666"
/translation="GSPAGRSIYNSFYVYCKPCQVOPGRLRVQSGTCROATLLTQ
GPSCHMDVLIIPNRMSGECOSPHCGTSAEFPKCGAPTSGKTSVALHLATSRNI
TCITCTDVRSPLVIFQNSRHVICLDGFHLVCTRLNDROVHDPOLGYSPLCV"

ORIGIN
Query Match 100.0%; Score 471; DB 6; Length 471;
Best Local Similarity 100.0%; Pred. No. 2,3e-143;
Matches 471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGAAGTCCAGCAGGTAGATCAATCTTACACAGCTTTTATGTGTGCAAGGCCCTGT 60
|||||
DB 1 GGAAGTCCAGCAGGTAGATCAATCTTACACAGCTTTTATGTGTGCAAGGCCCTGT 60
|||||
OY 61 CAAGAGTGCAGCGGGAACACTCAGGTACAGTGCAGACCTGCAGCAGGCAACGCTC 120

Db 61 CAAAGAGTGCAGCCGGAAAACTCAGGGTACAGTGCAGACCTGCAGAGCGCAACCTC 120
Qy 121 ACCGTGACCCAGGGTCCATCTTGCTGGAGATGATTTTAAATCCAAACCGATGAGTGT 180
Db 121 ACCGTGACCCAGGGTCCATCTTGCTGGAGATGATTTTAAATCCAAACCGATGAGTGT 180
Qy 181 GAATGCCAATCCCGACAGCTGCCCTGGAGCTAGTGCAGAAATTTTCTTTAAATGTGAGCA 240
Db 181 GAATGCCAATCCCGACAGCTGCCCTGGAGCTAGTGCAGAAATTTTCTTTAAATGTGAGCA 240
Qy 241 CACCCCACTCTGACAAAGAAACATCAGTACCTTTGACCTGATGCAACAATAGTCCG 300
Db 241 CACCCCACTCTGACAAAGAAACATCAGTACCTTTGACCTGATGCAACAATAGTCCG 300
Qy 301 AACATCATTGATAGTGCAGACAGTGCAGAGCCCGCTCGTCTGTTTCCAGTGCAC 360
Db 301 AACATCATTGATAGTGCAGACAGTGCAGAGCCCGCTCGTCTGTTTCCAGTGCAC 360
Qy 361 TCCCGCAGTGATTTGCTTAGACTGTTTCCACTTACTGTGTGACAAAGACTCAATGAT 420
Db 361 TCCCGCAGTGATTTGCTTAGACTGTTTCCACTTACTGTGTGACAAAGACTCAATGAT 420
Qy 421 CGGCACTTTGTTACGACACCTGCACTTGCTACTCCCTGCTTGTGTGTAG 471
Db 421 CGGCACTTTGTTACGACACCTGCACTTGCTACTCCCTGCTTGTGTGTAG 471

RESULT 2
BD269536 2960 bp DNA linear PAT 17-JUL-2003
LOCUS Mutations of the parkin gene, compositions, methods and uses.
DEFINITION BD269536
ACCESSION BD269536.1 GI:33079304
VERSION JP 2002536961-A/1.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2960)
AUTHORS Brice, A., Lucking, C., Abbas, N.E., Deneffe, P., Ricard, S. and Bouley, S.
TITLE Mutations of the parkin gene, compositions, methods and uses
JOURNAL AVENTIS PHARMA SA, INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE
COMMENT OS Homo sapiens (human)
PN JP 2002536961-A/1
PD 05-NOV-2002
PF 18-NOV-1999 JP 2000584062
PR 19-NOV-1998 FR 98/14524, 12-MAR-1999 US 60/124239 PR
PI 04-AUG-1999 FR 99/10140
PI ALEXIS BRICE, CHRISTOPHE LUCKING, NANCER EDDINE ABBAS, PATRICE PI DENEFFE,
PI SYLVAIN RICARD, SANDRINE BOULEY
PC C12N1/09, A01K67/027, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12Q1/68, G01N33/15, G01N33/50, C12N5/00, C12N5/00 CC
Mutations of the parkin gene, compositions, methods and uses FH
KEY Location/Qualifiers
FT source 1..2960
FT Location/Qualifiers
FT /organism='Homo sapiens (human)'.
1..2960
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN
Query Match 99.4%; Score 468; DB 6; Length 2960;
Best Local Similarity 100.0%; Pred. No. 2.9e-142;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAATCCAGAGGTGATCAATCTACAAACAGCTTTTATGTGATTCGAAAGCCCTGT 60
Db 504 GGAATCCAGAGGTGATCAATCTACAAACAGCTTTTATGTGATTCGAAAGCCCTGT 563
Qy 61 CAAAGAGTGCAGCCGGAAAACTCAGGGTACAGTGCAGACCTGCAGAGCGCAACCTC 120
Db 564 CAAAGAGTGCAGCCGGAAAACTCAGGGTACAGTGCAGACCTGCAGAGCGCAACCTC 623
Qy 121 ACCTTGACCCAGGGTCCATCTTGCTGGAGATGATTTTAAATCCAAACCGATGAGTGT 180
Db 121 ACCTTGACCCAGGGTCCATCTTGCTGGAGATGATTTTAAATCCAAACCGATGAGTGT 683
Qy 181 GAATGCCAATCCCGACAGCTGCCCTGGAGCTAGTGCAGAAATTTTCTTTAAATGTGAGCA 240
Db 181 GAATGCCAATCCCGACAGCTGCCCTGGAGCTAGTGCAGAAATTTTCTTTAAATGTGAGCA 240
Qy 241 CACCCCACTCTGACAAAGAAACATCAGTACCTTTGACCTGATGCCCAACAATAGTCCG 300
Db 241 CACCCCACTCTGACAAAGAAACATCAGTACCTTTGACCTGATGCCCAACAATAGTCCG 803
Qy 301 AACATCATTGATAGTGCAGACAGTGCAGAGCCCGCTCGTCTGTTTCCAGTGCAC 360
Db 301 AACATCATTGATAGTGCAGACAGTGCAGAGCCCGCTCGTCTGTTTCCAGTGCAC 863
Qy 361 TCCCGCAGTGATTTGCTTAGACTGTTTCCACTTACTGTGTGACAAAGACTCAATGAT 420
Db 361 TCCCGCAGTGATTTGCTTAGACTGTTTCCACTTACTGTGTGACAAAGACTCAATGAT 420
Qy 421 CGGCACTTTGTTACGACACCTGCACTTGCTACTCCCTGCTTGTGTGTAG 468
Db 421 CGGCACTTTGTTACGACACCTGCACTTGCTACTCCCTGCTTGTGTGTAG 971

RESULT 3
AF381283 904 bp mRNA linear PRI 02-MAY-2002
LOCUS Homo sapiens parkin isoform mRNA, complete cds, alternatively
DEFINITION AF381283
ACCESSION AF381283
VERSION AF381283.1 GI:20385799
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 904)
AUTHORS D'Agata, V., Scapagnini, G. and Cavallaro, S.
TITLE Functional and molecular diversity of parkin
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 904)
AUTHORS D'Agata, V., Scapagnini, G. and Cavallaro, S.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-2001) Institute of Biomedicine and Pathophysiology
of the Central Nervous System, Italian National Research Council
(CNR), Piazza Roma, 2, Catania 95123, Italy
Location/Qualifiers
source 1..904
/organism='Homo sapiens'
/mol_type='mRNA'
/db_xref='taxon:9606'
74..730
/note='alternatively spliced; transcript variant 7'
/codon_start=1
/product='parkin isoform'
/protein_id='P48214.8.1'
/db_xref='GI:20385800'
/translation='MNAATGGDDPRNAAAGCERBPQSLTRVDSLSSVLPQDSVGLAVIL
HTDSRKDSPAGSPAGRSIYNSFYVCKGCPQVQKLRVQCSCTROATLITLQGPS
CMBDVLIPNMSGECOSPHCPGSAEPFGCAHPISDKETPVALLILATNSNITCI
TCDIVRSPVLVFCNSRHVILCDPFLHYCTRLNDRQFVHDPQLGSLPCVVCLLRGM'

CDS
1..730
/note='alternatively spliced; transcript variant 7'
/codon_start=1
/product='parkin isoform'
/protein_id='P48214.8.1'
/db_xref='GI:20385800'
/translation='MNAATGGDDPRNAAAGCERBPQSLTRVDSLSSVLPQDSVGLAVIL
HTDSRKDSPAGSPAGRSIYNSFYVCKGCPQVQKLRVQCSCTROATLITLQGPS
CMBDVLIPNMSGECOSPHCPGSAEPFGCAHPISDKETPVALLILATNSNITCI
TCDIVRSPVLVFCNSRHVILCDPFLHYCTRLNDRQFVHDPQLGSLPCVVCLLRGM'

ORIGIN

Query Match 99.0%; Score 466.4; DB 9; Length 904;
Best Local Similarity 99.8%; Pred. No. 8.2e-142;
Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGTCACAGAGTATGATCAATCTACAAAGCTTTTATGTGTATTTGTAAGGCCCCCTGT 60
DB 239 GGAAGTCACAGAGTATGATCAATCTACAAAGCTTTTATGTGTATTTGTAAGGCCCCCTGT 298
QY 61 CAAAGATGACAGCGGGGAAAACTCAGGGTACAGTGCAGCACTGACAGGCAACGCTC 120
DB 299 CAAAGATGACAGCGGGGAAAACTCAGGGTACAGTGCAGCACTGACAGGCAACGCTC 358
QY 121 ACCTTGACCAAGAGTCCATCTTGTCTGGAGATGATTTTAAATCCAAACCGGATGATGCT 180
DB 359 ACCTTGACCAAGAGTCCATCTTGTCTGGAGATGATTTTAAATCCAAACCGGATGATGCT 418
QY 181 GAATGCCAATCCCACTGCGCTGGGACTAGTGCAGAAATTTTCTTAAATGTGAGCA 240
DB 419 GAATGCCAATCCCACTGCGCTGGGACTAGTGCAGAAATTTTCTTAAATGTGAGCA 478
QY 241 CACCCCACTCTGACAAAGAAACATCAGTACTTGGACCTGATCGCAACAAATAGTCGG 300
DB 479 CACCCCACTCTGACAAAGAAACATCAGTACTTGGACCTGATCGCAACAAATAGTCGG 538
QY 301 AACATCACTTGACATTAAGTGCAGACAGTGCAGAGCCCGCTGTGTTTCCAGTGCAC 360
DB 539 AACATCACTTGACATTAAGTGCAGACAGTGCAGAGCCCGCTGTGTTTCCAGTGCAC 598
QY 361 TCCCGCACAGTATTTGCTTATGACTGTTTCCATTAATCTGTGACAAAGATCAATGAT 420
DB 599 TCCCGCACAGTATTTGCTTATGACTGTTTCCATTAATCTGTGACAAAGATCAATGAT 658
QY 421 CGGAGTTTGTTCACGACCCCTCACTTGGCTACCTCCCTGCTGTGTG 468
DB 659 CGGAGTTTGTTCACGACCCCTCACTTGGCTACCTCCCTGCTGTGTG 706

RESULT 4
AR492172 2960 bp DNA linear PAT 15-MAY-2004
LOCUS AR492172
DEFINITION Sequence 1 from patent US 6716621.
ACCESSION AR492172
VERSION AR492172.1 GI:47260688
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2960)
AUTHORS Shimizu N. and Mizuno Y.
TITLE Isolated DNA or gene responsible for Parkinson's disease
JOURNAL Patent: US 6716621-A 1 06-APR-2004;
FEATURES
source 1..2960
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 99.0%; Score 466.4; DB 6; Length 2960;
Best Local Similarity 99.8%; Pred. No. 9.6e-142;
Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGTCACAGAGTATGATCAATCTACAAAGCTTTTATGTGTATTTGTAAGGCCCCCTGT 60
DB 504 GGAAGTCACAGAGTATGATCAATCTACAAAGCTTTTATGTGTATTTGTAAGGCCCCCTGT 563
QY 61 CAAAGATGACAGCGGGGAAAACTCAGGGTACAGTGCAGCACTGACAGGCAACGCTC 120
DB 564 CAAAGATGACAGCGGGGAAAACTCAGGGTACAGTGCAGCACTGACAGGCAACGCTC 623
QY 121 ACCTTGACCAAGAGTCCATCTTGTCTGGAGATGATTTTAAATCCAAACCGGATGATGCT 180
DB 624 ACCTTGACCAAGAGTCCATCTTGTCTGGAGATGATTTTAAATCCAAACCGGATGATGCT 683

QY 181 GAATGCCAATCCCACTGCGCTGGGACTAGTGCAGAAATTTTCTTAAATGTGAGCA 240
DB 684 GAATGCCAATCCCACTGCGCTGGGACTAGTGCAGAAATTTTCTTAAATGTGAGCA 743
QY 241 CACCCCACTCTGACAAAGAAACATCAGTACTTGGACCTGATCGCAACAAATAGTCGG 300
DB 744 CACCCCACTCTGACAAAGAAACATCAGTACTTGGACCTGATCGCAACAAATAGTCGG 803
QY 301 AACATCACTTGACATTAAGTGCAGACAGTGCAGAGCCCGCTGTGTTTCCAGTGCAC 360
DB 804 AACATCACTTGACATTAAGTGCAGACAGTGCAGAGCCCGCTGTGTTTCCAGTGCAC 863
QY 361 TCCCGCACAGTATTTGCTTATGACTGTTTCCATTAATCTGTGACAAAGATCAATGAT 420
DB 864 TCCCGCACAGTATTTGCTTATGACTGTTTCCATTAATCTGTGACAAAGATCAATGAT 923
QY 421 CGGAGTTTGTTCACGACCCCTCACTTGGCTACCTCCCTGCTGTGTG 468
DB 924 CGGAGTTTGTTCACGACCCCTCACTTGGCTACCTCCCTGCTGTGTG 971

RESULT 5
AX584224 2960 bp DNA linear PAT 10-JAN-2003
LOCUS AX584224
DEFINITION Sequence 1 from Patent WO02079459.
ACCESSION AX584224
VERSION AX584224.1 GI:27655742
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Jensen, P. H.
TITLE Means for inhibiting proteolytical processing of parkin
JOURNAL Patent: WO 02079459-A 1 10-OCT-2002;
FEATURES
source 1..2960
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
mRNA 1..2960

Query Match 99.0%; Score 466.4; DB 6; Length 2960;
Best Local Similarity 99.8%; Pred. No. 9.6e-142;
Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGTCACAGAGTATGATCAATCTACAAAGCTTTTATGTGTATTTGTAAGGCCCCCTGT 60
DB 504 GGAAGTCACAGAGTATGATCAATCTACAAAGCTTTTATGTGTATTTGTAAGGCCCCCTGT 563
QY 61 CAAAGATGACAGCGGGGAAAACTCAGGGTACAGTGCAGCACTGACAGGCAACGCTC 120
DB 564 CAAAGATGACAGCGGGGAAAACTCAGGGTACAGTGCAGCACTGACAGGCAACGCTC 623
QY 121 ACCTTGACCAAGAGTCCATCTTGTCTGGAGATGATTTTAAATCCAAACCGGATGATGCT 180
DB 624 ACCTTGACCAAGAGTCCATCTTGTCTGGAGATGATTTTAAATCCAAACCGGATGATGCT 683
QY 181 GAATGCCAATCCCACTGCGCTGGGACTAGTGCAGAAATTTTCTTAAATGTGAGCA 240
DB 684 GAATGCCAATCCCACTGCGCTGGGACTAGTGCAGAAATTTTCTTAAATGTGAGCA 743
QY 241 CACCCCACTCTGACAAAGAAACATCAGTACTTGGACCTGATCGCAACAAATAGTCGG 300
DB 744 CACCCCACTCTGACAAAGAAACATCAGTACTTGGACCTGATCGCAACAAATAGTCGG 803
QY 301 AACATCACTTGACATTAAGTGCAGACAGTGCAGAGCCCGCTGTGTTTCCAGTGCAC 360
DB 804 AACATCACTTGACATTAAGTGCAGACAGTGCAGAGCCCGCTGTGTTTCCAGTGCAC 863

| | | | | |
|---------------------------|--|---------------------|--|------------------------|
| Oy | | 361 | TCCGCCACGCGATTTCCTTAAGACTGTGGTTTCCACTTAACTACGTGTGCACAGACTCAATGAT | 420 |
| Db | | 864 | TCCGCCACGCGATTTCCTTAAGACTGTGGTTTCCACTTAACTACGTGTGCACAGACTCAATGAT | 923 |
| Oy | | 421 | CGGCAGTTTGTTTCACGACCCTCAACTTGGGCTAACCCTGCTGTGTGG | 468 |
| Db | | 924 | CGGCAGTTTGTTTCACGACCCTCAACTTGGGCTAACCCTGCTGTGTGG | 971 |
| RESULT 6 AX584226 | | | | |
| LOCUS | AX584226 | 2960 bp | DNA | linear PAT 10-JUN-2002 |
| DEFINITION | Sequence 3 from Patent WO02079459. | | | |
| ACCESSION | AX584226 | | | |
| VERSION | AX584226.1 | GI:27655743 | | |
| KEYWORDS | | | | |
| SOURCE | Homo sapiens (human) | | | |
| ORGANISM | Homo sapiens | | | |
| REFERENCE | Jensen, P.H. Means for inhibiting proteolytical processing of parkin | | | |
| AUTHORS | Patent: WO 02079459-A 3 10-OCT-2002; | | | |
| JOURNAL | NSGENE A/S (DK) | | | |
| FEATURES | Location/Qualifiers | | | |
| source | 1..2960 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" | | | |
| ORIGIN | mRNA 1..2960 | | | |
| Query Match | 99.0%; | Score 466.4; | DB 6; | Length 2960; |
| Best Local Similarity | 99.8%; | Pred. No. 9.6e-142; | | |
| Matches 467; Conservative | 0; | Mismatches 1; | Indels 0; | Gaps 0; |
| Oy | | 1 | GGNAGTCGCGCGGTGATCAATCTTACAACAGCTTTTATGTGTATTTGCCAAGGCCCTGT | 60 |
| Db | | 504 | GGAAGTCGCGCGGTGATCAATCTTACAACAGCTTTTATGTGTATTTGCCAAGGCCCTGT | 563 |
| Oy | | 61 | CAAGAAGTCAGCCGGGAAAACCTCAGGGGTACAGTGCAGACCTGCAGGCAAGGCAACGCTC | 120 |
| Db | | 564 | CAAAAGTGCAGCCGGGAAAACCTCAGGGGTACAGTGCAGACCTGCAGGCAAGGCAACGCTC | 623 |
| Oy | | 121 | ACCTTGACCCAGGGTCCATCTTGTCTGGGATGATTTTTTAATCCAAAACCGATAGTGT | 180 |
| Db | | 624 | ACCTTGACCCAGGGTCCATCTTGTCTGGGATGATTTTTTAATCCAAAACCGATAGTGT | 683 |
| Oy | | 181 | GAATGCCAATCCCCACAATGCCCTGGGACCTAGTGCAGAAATTTTCTTTAATGTGAGCA | 240 |
| Db | | 684 | GAATGCCAATCCCCACAATGCCCTGGGACCTAGTGCAGAAATTTTCTTTAATGTGAGCA | 743 |
| Oy | | 241 | CACCCGACCTGTGCAGAGAACAATCATGAGCTTTGACCTGATCGCAACAAATATGTCG | 300 |
| Db | | 744 | CACCCGACCTGTGCAGAGAACAATCATGAGCTTTGACCTGATCGCAACAAATATGTCG | 803 |
| Oy | | 301 | AACATCACCTTGATTAAGTGCACAGAGTCAAGAGCCCGCTCGTGTTTCCAGTCAAC | 360 |
| Db | | 804 | AACATCACCTTGATTAAGTGCACAGAGTCAAGAGCCCGCTCGTGTTTCCAGTCAAC | 863 |
| Oy | | 361 | TCCGCCACGCGATTTCCTTAAGACTGTGGTTTCCACTTAACTACGTGTGCACAGACTCAATGAT | 420 |
| Db | | 864 | TCCGCCACGCGATTTCCTTAAGACTGTGGTTTCCACTTAACTACGTGTGCACAGACTCAATGAT | 923 |
| Oy | | 421 | CGGCAGTTTGTTTCACGACCCTCAACTTGGGCTAACCCTGCTGTGTGG | 468 |
| Db | | 924 | CGGCAGTTTGTTTCACGACCCTCAACTTGGGCTAACCCTGCTGTGTGG | 971 |
| RESULT 7 AB009973 | | | | |
| LOCUS | AB009973 | 2960 bp | mRNA | linear PRI 14-APR-2000 |

| | |
|------------|---|
| DEFINITION | Homo sapiens mRNA for Parkin, complete cds. |
| ACCESSION | AB009973 |
| VERSION | AB009973.1 GI:3063387 |
| KEYWORDS | Parkin. |
| SOURCE | Homo sapiens (human) |
| ORGANISM | Homo sapiens |
| REFERENCE | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. |
| AUTHORS | 1 (sites) |
| TITLE | Kitada,T., Asakawa,S., Hattori,N., Matsumine,H., Yamamura,Y., Minoshima,S., Yokochi,M., Mizuno,Y. and Shimizu,N. |
| JOURNAL | Mutations in the parkin gene cause autosomal recessive juvenile Parkinsonism |
| MEDLINE | Nature 392 (6676), 605-608 (1998) |
| PUBMED | 98219084 |
| REFERENCE | 9560156 |
| AUTHORS | 2 (bases 1 to 2960) |
| TITLE | Shimizu,N., Kitada,T. and Asakawa,S. |
| JOURNAL | Direct Submision |
| DEFINITION | Submitted (22-DEC-1997) Nobuyoshi Shimizu, Keio University School of Medicine, Department of Molecular Biology, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan |
| DEFINITION | (E-mail:shimizu@med.keio.ac.jp, Tel:81-3-3351-2370(ex.2720), Fax:81-3-3351-2370) |
| FEATURES | Location/Qualifiers |
| source | 1..2960 |
| gene | /organism="Homo sapiens" |
| gene | /mol_type="mRNA" |
| gene | /db_xref="taxon:9606" |
| gene | /chromosome="6" |
| gene | /map="6q25.2-q27" |
| gene | 1..2960 |
| gene | /gene="Parkin" |
| gene | 1..108 |
| gene | /gene="Parkin" |
| gene | /number=1 |
| gene | 102..1499 |
| gene | /gene="Parkin" |
| gene | /codon_start=1 |
| gene | /product="Parkin" |
| gene | /protein_id="PAA25751.1" |
| gene | /db_xref="GI:3063388" |
| gene | /translation="MTVFRFNSHGFPEVENSIDSITFOLKEVAKRGQVPADQLRTV PAGELNDMTVQNCDDQSIIVHIVQFWRKQEQENWATGDDPNNAGCGEREPOSIT TRVLISSVLPDGSVGLAVTLHTDSRKDPSPAGSEIYNSFYVCKGQCORPOPC KLRFQCTCQATQTLTLTGSPSCMDVLIENRSGCQSPHCSTSAEEFFKGAAPT DKRPPVAILIATNSRNITCICTCTDVRSVILVFOCNSRVICLDGFHLVCTPLNDRC FVHPDQGYSLPCVAGCPNSLIKEIHPRILGEEQYNYOOGAECVLCMGCVLCSP FGAGGLLPEDDQKVTCEGNGLCGAPACRCEKAEYHBSGCSAVFASGTTQATRT VDEBAQAQARWEASKEITIKTTTPCPRCHAVEVERKNGCQMHKCPQCRLEMCNCG CEWRVMDGHWFDV" |
| exon | 109..272 |
| exon | /gene="Parkin" |
| exon | /number=2 |
| exon | 273..513 |
| exon | /gene="Parkin" |
| exon | /number=3 |
| exon | 514..635 |
| exon | /gene="Parkin" |
| exon | /number=4 |
| exon | 636..719 |
| exon | /gene="Parkin" |
| exon | /number=5 |
| exon | 720..835 |
| exon | /gene="Parkin" |
| exon | /number=6 |
| exon | 836..972 |
| exon | /gene="Parkin" |
| exon | /number=7 |
| exon | 973..1034 |
| exon | /gene="Parkin" |
| exon | /number=8 |
| exon | 1035..1184 |

| | | | |
|--|-----------------|---|--------------------|
| | | /gene="parkin" | |
| exon | | /number=9 1185..1268 | |
| exon | | /gene="parkin" /number=10 1269..1386 | |
| exon | | /gene="parkin" /number=11 1387..2960 | |
| exon | | /gene="parkin" /number=12 | |
| ORIGIN | | | |
| Query Match | 99.0%; | Score 466.4; | DB 9; Length 2960; |
| Best Local Similarity | 99.8%; | Pred. No. 9.6e-142; | |
| Matches 467; | Conservative 0; | Mismatches 1; | Indels 0; Gaps 0; |
| OY | 1 | GGAAGTCGACGGATGCATCTCAACAAGCCTTTTATGTATATTGGAAAGGCCCGTG | 60 |
| Db | 504 | GGAGTCGACGAGGATGCATCTCAACAAGCCTTTATGTATATTGGAAAGGCCCGTG | 563 |
| OY | 61 | CAGAAGTCGACGCCGGAAAATCTCAGGGTACAGTCAGCACCTCCGACGACGCAAGCTC | 120 |
| Db | 564 | CAAAGATGACGACC GGAAAACTCAGGGTACAGTCAGCACCTCCGACGACGCAAGCTC | 623 |
| OY | 121 | AACCTTGACCCAGGGTCCACTCTTGTGGATGATGTTTTAATTCAAACCGCATGAGTGT | 180 |
| Db | 624 | ACCTTGACCCAGGGTCCACTCTTGTGGATGATGTTTTAATTCAAACCGCATGAGTGT | 683 |
| OY | 181 | GAATGCCAATCCCCACACTGCCCCTGGGACTAGTCAGAAATTTTCTTTAAATGTGGACA | 240 |
| Db | 684 | GAATGCCAATCCCCACACTGCCCCTGGGACTAGTCAGAAATTTTCTTTAAATGTGGACA | 743 |
| OY | 241 | CACCCACCTCTGACAAAGAAAACATCATGAGCTTTGCACCTGATCGCAAAATATAGTGG | 300 |
| Db | 744 | CACCCACCTCTGACAAAGAAAACATCATGAGCTTTGCACCTGATCGCAAAATATAGTGG | 803 |
| OY | 301 | AACATCACTTGATTAAGSTGACAGACGTCAGAGACCCCGTCTGTTTCCAGTGCAAC | 360 |
| Db | 804 | AACATCACTTGATTAAGSTGACAGACGTCAGAGACCCCGTCTGTTTCCAGTGCAAC | 863 |
| OY | 361 | TCCGCCACAGTGATTTGCTTAGACTGTTTCCCTTAATCTGTGTGACAAGACTCAATAT | 420 |
| Db | 864 | TCCGCCACAGTGATTTGCTTAGACTGTTTCCCTTAATCTGTGTGACAAGACTCAATAT | 923 |
| OY | 421 | CGGCGATTGTTACAGCACCCCTCAACTTGGACTACCTCCGCTGTGTG | 468 |
| Db | 924 | CGGCGATTGTTACAGCACCCCTCAACTTGGACTACCTCCGCTGTGTG | 971 |
| RESULT 8 | | 1575 bp mRNA linear PRI 06-OCT-2001 | |
| LOCUS BC022014 | | | |
| DEFINITION Homo sapiens Parkinson disease (autosomal recessive, juvenile) 2, | | | |
| LOCUS LINKAGE Parkin, mRNA (CDNA clone MGC:26491 IMAGE:4824892), complete cds. | | | |
| ACCESSION BC022014 | | | |
| VERSION BC022014.2 GI:34191069 | | | |
| KEYWORDS MGC. | | | |
| SOURCE Homo sapiens (human) | | | |
| ORGANISM Homo sapiens | | | |
| REFERENCE Euteryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 1575) Klausner,R.D., Collins,F.S., Wagner,L., Sherman,C.M., Schuler,G.D., Altehusen,S.F., Zeeberg,B., Bueter,K.H., Scheffer,C.P., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stjepanec,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheer,T.E., Brownstein,M.J., Ueda,T.B., Tohtiyuki,S., Carninci,P., Prange,C., Raha,S.S., Logueliano,N.A., Peters,G.J., Abrahamson,R.D., Mullany,S.J., Bosak,S.A., McSwain,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., | | | |

| | |
|--------------|--|
| TITLE | Journal of Molecular Biology |
| JOURNAL | Journal of Molecular Biology |
| MEDLINE | 22388927 |
| PUBMED | 12477932 |
| REFERENCE | 2 (bases 1 to 1575) |
| AUTHORS | Strauberg, R. |
| TITLE | Direct Submision |
| JOURNAL | Submitted (22-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA |
| REMARK | NIH-MGC Project URL: http://mgc.nci.nih.gov On Aug 25, 2003 this sequence version replaced gi:18314632. Contact: MGC help desk Email: gcgaps-remail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN) CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Institute for Systems Biology http://www.systemsbioology.org contact: amadansystemsbiology.org Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whitting |
| FEATURES | Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Series: IRAC Plate: 33 Row: 1 Column: 6 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4758883. Location/Qualifiers 1..1575 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:26491 IMAGE:4824892" /tissue_type="Testis" /clone_1lb="NTH MGC_97" /lab_host="DH10B" /note="Vector: pBluescript" |
| gene | 1..1575 /gene="PARK2" /note="synonyms: AR-JP, PDJ, PRKN" /db_xref="locusid:5071" /db_xref="MIM:602544" 99..1262 /gene="PARK2" /codon_start=1 /product="PARK2 protein" /protein_id="AAH22014.1" /db_xref="GI:18314633" /db_xref="LocusID:5071" /db_xref="MIM:602544" /reanalysis="M1VFVFNSSHGPFVEVDSTSIIFOLKEVAKRGVADOLRVV PAGEHLENDWTVNCGDLDOOSIYHIVORPMAGGEMATGDDPRNMAAGGGERE TRVDSLSVVPDSSVGLAVLHMISRVSPRSPKPARSIYNSPVCYGCORVOP KRVQCTCROATTTTQGRSCMDVLIIPRMSECSPHCRGSAEPFPKCAHPT DKETSVLALIALNSRNITCTIDVRSPLVIFQNSRHVILCQFHLVCTRLNDR FVHPDQAGSLPCVGTGDTVLRGALGFRGVAGCPNSLKEHLHFFILAEEDYNR QQYAEAEICLQMGVCLPRBCGAGLLPEPDORVTCGEGNGLCGYGORRTK" 99..317 /gene="PARK2" /note="ubiquitin; Region: Ubiquitin family. This family contains a number of ubiquitin-like proteins: SUMO (smc3 |
| CDS | |
| misc_feature | |

ORIGIN
homologue), Ned8, Elongin B, Rub1"
/db_xref="CDD:pfam00240"

Query Match 98.7%; Score 464.8; DB 9; Length 1575;
Best Local Similarity 99.6%; Pred. No. 3e-141;
Matches 466; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGTCAGAGGTAGTCAATCTACAGAGCTTTTATGTATTCGAAAGCCCTCTG 60
DB 501 GGAAGTCAGAGGTAGTCAATCTACAGAGCTTTTATGTATTCGAAAGCCCTCTG 560
QY 61 CAAAGAGTCAGAGCCGGGAAAACCTCAGGGTACAGTGAGCAGCCTGACGAGGAGCGCTC 120
DB 561 CAAAGAGTCAGAGCCGGGAAAACCTCAGGGTACAGTGAGCAGCCTGACGAGGAGCGCTC 620
QY 121 ACCCTGACCGAGGGTCCATCTTGTCTGGATGATGTTTAAATTCGAAACCGGATAGTGGT 180
DB 621 ACCCTGACCGAGGGTCCATCTTGTCTGGATGATGTTTAAATTCGAAACCGGATAGTGGT 680
QY 181 GAATGCCAATCCCACTGCTCCCTGGGACTAGTGAGCAAAATTTTCTTTAAATGTGGAGA 240
DB 681 GAATGCCAATCCCACTGCTCCCTGGGACTAGTGAGCAAAATTTTCTTTAAATGTGGAGA 740
QY 241 CACCCCACTCTGACAGAGAAACATAGTACTTTGACCTGATGCAACAAATAGTGG 300
DB 741 CACCCCACTCTGACAGAGAAACATAGTACTTTGACCTGATGCAACAAATAGTGG 800
QY 301 AACATCACTTGATTCGTGCAAGAGCGTCAAGAGCCCGTCTGCTTTTCCAGTGCAC 360
DB 801 AACATCACTTGATTCGTGCAAGAGCGTCAAGAGCCCGTCTGCTTTTCCAGTGCAC 860
QY 361 TCCCGCAGCTGATTTGCTTAGACTGTTTCACTTATCTGTGACAAGACTCAATGAT 420
DB 861 TCCCGCAGCTGATTTGCTTAGACTGTTTCACTTATCTGTGACAAGACTCAATGAT 920
QY 421 CGGCACTTTGTCAGAGCCCTCACTGCTACTCCCTGCTTGTGTG 468
DB 921 CGGCACTTTGTCAGAGCCCTCACTGCTACTCCCTGCTTGTGTG 968

RESULT 9
AF381282 1157 bp mRNA linear PRI 02-MAY-2002
LOCUS Homo sapiens parkin isoform mRNA, complete cds, alternatively
DEFINITION AF381282
ACCESSION AF381282
VERSION AF381282.1 GI:20385797
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1157)
D'Agata,V., Scapagnini,G. and Cavallaro,S.
JOURNAL Functional and molecular diversity of parkin
AUTHORS Unpublished
TITLE 2 (bases 1 to 1157)
JOURNAL D'Agata,V., Scapagnini,G. and Cavallaro,S.
AUTHORS Direct Submission
TITLE Submitted (01-MAY-2001) Institute of Bioimaging and Pathophysiology
JOURNAL of the Central Nervous System, Italian National Research Council
REFERENCE (CNR), Piazza Roma, 2, Catania 95123, Italy
FEATURES
SOURCE Location/Qualifiers
1..1157
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
333..1157
/note="alternatively spliced; transcript variant 6"
/codon_start=1
/product="parkin isoform"
/protein_id="AAW21457.1"

/db_xref="GI:20385798"
/translation="MSGECSPHCPGTSAEFFKCGAAPTSDKETPVALLIATNSRN
ITCITCDVRSPLVLPQNSRHVI CLDFCHLYCVTRINDPQFVDPOLGVSLPVAC
PNSLIKEIHPRIIIEGYNRYOQYAGAEFCVILQMGVLCRPGGALLIEPPDRKTT
CEGNGAGCCGAPRBECKEAYHEBSCAVRBSGTTTQATRVDRALAAEQARWENASKE
TIKTYTRCPCHVPEKNGGCMKCPQPCRLIEWCMNGCEWNRVCMGDHWFV"

ORIGIN
Query Match 97.8%; Score 460.8; DB 9; Length 1157;
Best Local Similarity 99.6%; Pred. No. 5.9e-140;
Matches 462; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GTCCAGCAGGTAGATCAATCTACAGAGCTTTTATGTATTCGAAAGCCCTGTCAA 64
DB 166 GTCCAGCAGGTAGATCAATCTACAGAGCTTTTATGTATTCGAAAGCCCTGTCAA 225
QY 65 GAGTGACGCGGAAAACCTCAGGGTACAGTGAGCAGCCTGACGAGGAGCGCTACCT 124
DB 226 GAGTGACGCGGAAAACCTCAGGGTACAGTGAGCAGCCTGACGAGGAGCGCTACCT 285
QY 125 TGACCCAGGGTCCATCTTGTCTGGATGATGTTTAAATTCGAAACCGGATAGTGGAA 184
DB 286 TGACCCAGGGTCCATCTTGTCTGGATGATGTTTAAATTCGAAACCGGATAGTGGAA 345
QY 185 GCCAATCCCACTGCTCCCTGGGACTAGTGCAAAATTTTCTTTAAATGTGGAGCAC 244
DB 346 GCCAATCCCACTGCTCCCTGGGACTAGTGCAAAATTTTCTTTAAATGTGGAGCAC 405
QY 245 CCACCTCTGACAGAAACATCAGTACTTTGACCTGATGCAACAAATAGTGGAGACA 304
DB 406 CCACCTCTGACAGAAACATCAGTACTTTGACCTGATGCAACAAATAGTGGAGACA 465
QY 305 TCACCTGATTAAGTGACAGAGCTGAGAGCCCGTCTGTTTCCAGTCAACTCC 364
DB 466 TCACCTGATTAAGTGACAGAGCTGAGAGCCCGTCTGTTTCCAGTCAACTCC 525
QY 365 GCCACGTGATTTGCTTAGACTGTTTCACTTATCTGTGACAAGACTCAATGATGGG 424
DB 526 GCCACGTGATTTGCTTAGACTGTTTCACTTATCTGTGACAAGACTCAATGATGGG 585
QY 425 AGTTTGTTCAGAGCCCTCACTGCTACTCCCTGCTTGTGTG 468
DB 586 AGTTTGTTCAGAGCCCTCACTGCTACTCCCTGCTTGTGTG 629

RESULT 10
AB060701 583 bp mRNA linear MAM 20-APR-2002
LOCUS Bos taurus mRNA for parkin, partial cds.
ACCESSION AB060701
VERSION AB060701.1 GI:16444927
KEYWORDS
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
REFERENCE 1
Kitada,T., Asakawa,S., Hattori,N., Matsumine,H., Yamamura,Y.,
Mitsushima,S., Yokochi,M., Mizuno,Y. and Shimizu,T.
TITLE Mutations in the parkin gene cause autosomal recessive juvenile
JOURNAL parkinsonism
AUTHORS Nature 392 (6676), 605-608 (1998)
MEDLINE 98219084
PUBMED 9560156
REFERENCE 2
Haase,A., Yamada,H., Arai,K., Sunada,Y., Shimizu,T. and Matsumura,K.
TITLE Characterization of parkin in bovine peripheral nerve
JOURNAL Brain Res. 930 (1-2), 143-149 (2002)
MEDLINE 21869006
PUBMED 11879804
REFERENCE 3 (bases 1 to 583)
Haase,A., Yamada,H., Arai,K., Shimizu,T. and Matsumura,K.

TITLE Direct Submission
JOURNAL Submitted (25-APR-2001) Asako Hase, Teikyo University School of
Medicine, Dept. of Neurology and Neuroscience; 2-11-1, Kaga,
Itabashi-ku, Tokyo 173-8605, Japan (E-mail:hase@med.teikyo-u.ac.jp,
Tel:81-3-3964-1211 (ex.1916), Fax:81-3-3964-6394)

FEATURES
source
1. .583
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="brain"
/note="Identical sequence is also seen in peripheral
nerve"

gene
1. .583
/gene="parkin gene"
/db_xref="GI:16444928"
/protein_id="BAB70670.1"
/evidence="experimental"
/product="parkin"
/translation="AGCEREPOSLETRVDSLSSVLPEDSVGLAVILHTDSRKDPPAG
SPAGRSYNSFFVYCKGRCQVOPGKLRVOCSTQATLTLLTGSPSCMDVLTENRMS
SGCOSPHCPGSAEPFKCGAHPSTDEKTSVALHLIATNSRNITCITDVRSPVLVF
QNSRHVICDCHLYCVRILNDQPFVHDQLGY"

ORIGIN

Query Match 96.4%; Score 454; DB 4; Length 583;
Best Local Similarity 100.0%; Pred. No. 9.3e-138;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTCACAGAGTATGATCAATCTACACAGCTTTATGATGATTCGAAAGGCCCTGT 60
DB 130 GGAAGTCACAGAGTATGATCAATCTACACAGCTTTATGATGATTCGAAAGGCCCTGT 189
QY 61 CAAAGATGACAGCCGGGAAACTCAGGGTACAGTGCAGACCTGCAGCAGCAAGCCTC 120
DB 190 CAAAGATGACAGCCGGGAAACTCAGGGTACAGTGCAGACCTGCAGCAGCAAGCCTC 249
QY 121 ACCTTGAACCCAGGGTCAATCTTGGATGATGATGATGATGATGATGATGATGATGAT 180
DB 250 ACCTTGAACCCAGGGTCAATCTTGGATGATGATGATGATGATGATGATGATGATGAT 309
QY 181 GAATGCCAATCCCACTGCTGCTGGACTGAGTATGTCAGAAATTTCTTTAAATGTGAGCA 240
DB 310 GAATGCCAATCCCACTGCTGCTGGACTGAGTATGTCAGAAATTTCTTTAAATGTGAGCA 369
QY 241 CACCCCACTCTGACAGAAACATCAGTACCTTGACCTGATCGCAACAAATAGTCGG 300
DB 370 CACCCCACTCTGACAGAAACATCAGTACCTTGACCTGATCGCAACAAATAGTCGG 429
QY 301 AACATCACTTGATTAAGTGCAGACGTCAGAGCCCGCTCTGATTTCCAGTGCAC 360
DB 430 AACATCACTTGATTAAGTGCAGACGTCAGAGCCCGCTCTGATTTCCAGTGCAC 489
QY 361 TCCCGCCACGATTTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 490 TCCCGCCACGATTTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 549
QY 421 CCGCAGTTTGTTCAGACCCCTCAACTGGCTACT 454
DB 550 CCGCAGTTTGTTCAGACCCCTCAACTGGCTACT 583

RESULT 11
AB039878 1564 bp mRNA linear ROD 16-APR-2004
DEFINITION Rattus norvegicus mRNA for Parkin, complete cds.
ACCESSION AB039878
VERSION AB039878.1 GI:7229096
KEYWORDS Parkin.
SOURCE Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS
1 Hattori, N., Shimura, H., Kubo, S., Kitada, T., Wang, M., Asakawa, S.,
Minashima, S., Shimizu, N., Suzuki, T., Tanaka, K. and Mizuno, Y.
Autosomal recessive juvenile parkinsonism: A key to understanding
neural degeneration in sporadic Parkinson's disease
Neuropathology 20 81, 85-90 (2000)
2 (bases 1 to 1564)
Hattori, N., Wang, M. and Mizuno, Y.
Direct Submission
JOURNAL
Submitted (07-MAR-2000) Nobutaka Hattori, Junendo University,
Dept. of Neurology, Hongo 2-1-1, Bunkyo, Tokyo 113-8421, Japan
(E-mail:hattori@med.junendo.ac.jp, Tel:81-3-3813-3111 (ex.3320),
Fax:81-3-3813-7440)

FEATURES
source
1. .1564
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Srague-Dawley"
/db_xref="taxon:10116"
/sex="male"
/clone_id="recombinant bacterial type"
/dev_stage="adult"
/tissue_type="brain"
/note="vector:pAP3neo"

gene
1. .1564
/gene="parkin"
/db_xref="GI:7229097"
/translation="MIVFVRFNSYGFPEVDSPTSIFOLKEVAKRGVPAQDLAVI
FAGKELNHLTVONCDLEQOSIVHIVORPKSHETNAGDGKQSTPEGSIIPRSI
TRYDLSHLTPADSVGLAVILPDSKSDSRAAGRAKPTPHYSFFVCKGCHVOPG
KLVQCGCTQATLTLLAOGPSCMDVLIIPRMSECCSPDCRGTRAEFPKCAHPTS
DKDTSVALNLTNNRSRIPCIACDVNPVLPFCNRRHYICLDFHLYCVRILNDRO
FVHDAQVSLPCVAGCPNSI,IKELHFRILGEGVNRVYQYGAEBCLVLDQVGLCPK
PVGAGLIPROGKRTCEGANGAGCGFVRCRDKEAYHEGECDSMEASGATQAYR
VDRAAQRARMEASKETIKKTKPCRCNVPIEKNGCMHKRPOQCKLEWCNCG
CEMRACMGHMDV"

ORIGIN

Query Match 75.0%; Score 353.4; DB 10; Length 1564;
Best Local Similarity 84.8%; Pred. No. 1.4e-104;
Matches 396; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 2 GAATCCAGCAGTATGATCAATCTTACAAACGCTTTATGATGATTCGAAAGGCCCTGTC 61
DB 474 GAATCCAGCAGTATGATCAATCTTACAAACGCTTTATGATGATTCGAAAGGCCCTGTC 533
QY 62 AAAAGTGCAGCCGGGAAACTCAGGGTACAGTGCAGACCTGCAGGAGGCAACGCTCA 121
DB 534 AAAAGTGCAGCCGGGAAACTCAGGGTACAGTGCAGACCTGCAGGAGGCAACGCTCA 593
QY 594 CTTGGCCACAGGGCCCATCTTGGCTGGATGATGATGATGATGATGATGATGATGATGATGAT 653
DB 182 AATGCCAATCCCACTGCTGCTGGACTGATGATGATGATGATGATGATGATGATGATGATGAT 241
QY 122 CTTGACCCAGGGTCAATCTTGTGCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 181
DB 594 CTTGGCCACAGGGCCCATCTTGGCTGGATGATGATGATGATGATGATGATGATGATGATGAT 653
QY 654 AGTGCAATCTCCAGACTGCTGGCAAGAGCTGAATTTTCTTTAAATGTGAGAGAC 713
DB 242 ACCCACTCTGACAAAGAAACATCAGTACCTTGACCTGATGCAACAAATGTGCGGA 301
QY 714 ACCCACTCTGACAAAGAAACATCAGTACCTTGATGATGATGATGATGATGATGATGATGATGAT 773
DB 302 ACATCACTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 361

Db 774 GCATCCCTGCATCGCGTGACGAGATGCAGAACCCCTGCTTGTCTTCCATGTAACC 833
Qy 362 CCCGACGAGATTTGCTTGAAGCTTTTGCATTAATAGTGTGACAACTCAATGATC 421
Db 834 ACCGCGACGTGATCTTTGGACTGCTTGCATTTGTAAGTGTGCACAAAGGCTCAACGATC 893
Qy 422 GGCAGTTTGTTCACGACCTTCAACTTGGCTACTCCCTGCTTGTGTG 468
Db 894 GGCAGTTTGTTCACGACGCTCAACTTGGCTACTCCCTGCTTGTGTG 940

RESULT 12
AF343575 904 bp mRNA linear ROD 02-FEB-2002
LOCUS Rattus norvegicus parkin transcript variant 7 mRNA, complete cds.
DEFINITION AF343575
ACCESSION AF343575.1 GI:18478869
VERSION
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 904)
D'Agata,V.M., Scapagnini,G. and Cavallaro,S.
Functional and molecular diversity of parkin in the rat brain
Unpublished
2 (bases 1 to 904)
D'Agata,V.M., Scapagnini,G. and Cavallaro,S.
Direct Submission
Submitted (29-JAN-2001) Institute of Biomedicine and Pathophysiology
of the Central Nervous System, Italian National Research Council
(CNR), Piazza Roma, 2, Catania 95123, Italy
Location/Qualifiers
1..904
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
410..745
/codon_start=1
/product="parkin transcript variant 7"
/protein_id="AA173348.1"
/db_xref="GI:18478870"
/translation="MSGECQSPDCPGRARFFPKGAHPSDKDTSVALNLITNNSRS
IPCIACIDVRNPVLVFOCMNRHVI.CLDPCFHLVCTRLNDROFVADQLGSLPCVVRP
LPGLQGSIP"

ORIGIN
Query Match 74.4%; Score 350.2; DB 10; Length 904;
Best Local Similarity 84.4%; Pred. No. 1.5e-103;
Matches 394; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Db 2 GAAGTCCAGAGATGATCAATCTTCAACAGCTTTTATGTATGTGAAAGGCCCTGTC 61
Qy 240 GAGGTCCAGAGCTTAAACCTCACTACACAGCTTTTGTCTACTGCAAGGCCCTGTC 299
Db 62 AAAGATGACAGCCGCGAAACTCAGGTTACAGTCAAGTCAAGGCAAGGCAAGCTCA 121
Qy 300 ACAAGTTCAGCTCGGAACTCCGAGTTCAAGTGGGCACTGCAAGCAAGCAACCTCA 359
Db 122 CCTTGACCCAGGTCATCTTGTCTGGATGATGTTTAAATTCACAAACCGATGATGTG 181
Qy 360 CTTTGCGCCAGGCGCCATCTTGTCTGGATGATGTTTAAATTCACAAACCGATGATGTG 419
Db 182 AATGCCAATCCCACTGCGCTTGGGCTAGTGCAGAAATTTTCTTAAATGTGAGACAC 241
Qy 420 AGTGCCATCTCCAGATCTGCCCTTGGGCAAGAGCTGAATTTTCTTAAATGTGAGACAC 479
Db 242 ACCCCACTTGTGACAAAGAAACATCACTAGCTTTGCACTGATGCAACAAATAGTGTGA 301
Qy 480 ACCCAACTGACAAAGGACACATCACTAGCTTTGAACTGATCACCACAAAGCGCGCA 539
Db 302 ACATCACTTGATTAAGTGCACAGAGCTCAGAGCCCGGCTGCTTGTTCAGTGAACCT 361

Db 540 GCATCCCTGCATCGCGTGACGAGATGTCAGAAACCCCTGCTTGTCTTCCATGTAACC 599
Qy 362 CCCGACGAGATTTGCTTGAAGCTTTTGCATTAATAGTGTGACAAAGCTCAATGATC 421
Db 600 ACCGCGACGTGATCTTTGGACTGCTTGCATTTGTAAGTGTGCACAAAGGCTCAACGATC 659
Qy 422 GGCAGTTTGTTCACGACCTTCAACTTGGCTACTCCCTGCTTGTGTG 468
Db 660 GGCAGTTTGTTCACGACGCTCAACTTGGCTACTCCCTGCTTGTGTG 706

RESULT 13
AF168004 1466 bp mRNA linear ROD 19-FEB-2000
LOCUS Rattus norvegicus parkin mRNA, partial cds.
DEFINITION AF168004
ACCESSION AF168004.1 GI:7001383
VERSION
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1466)
D'Agata,V., Zhao,W. and Cavallaro,S.
Cloning and distribution of the rat parkin mRNA
Brain Res. Mol. Brain Res. 75 (2), 345-349 (2000)
PUBMED 10686358
2 (bases 1 to 1466)
D'Agata,V. and Cavallaro,S.
Direct Submission
Submitted (12-JUL-1999) Institute of Biomedicine and Pathophysiology
of the Central Nervous System, Italian National Research Council,
Piazza Roma, 2, Catania 95123, Italy
Location/Qualifiers
1..1466
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/tissue_type="brain"
91..>1466
/codon_start=1
/product="Parkin"
/protein_id="AA34874.1"
/db_xref="GI:7001384"
/translation="MIVFREFNSYGFPEVDSDTSLCQKEVAKRGQVPAQDLRLVI
FAGKELOHNLTVQNCDELOOSIVHIVORPORSKSHENASGDKPQSTPEGSIMPEPSL
TRVDLSHILPADSVGLAVILDPDTSKSDSEARGBAKPTYSFVYCGKPGKXQPG
KLRVCGCTGROATLTAOGPSCMDVLIENRMSGCQSPDCGTGAABFFPKGAHPS
DKDTSVALNLITNNSRSTPCTACTVRNPVLVFOCMNRHVI.CLDPCFHLVCTRLNDRO
FVADQLGSLPCVACGNSLKEIHLHFRITLDEBQINRIQQYGAEECVLQMGVLCPR
PGGAGLPEQGVQKRVCTEGAGLGGVPCFDCDEAHEGSDMFEASGATISQAYR
VDORABQARMEASKEITIKTKTKPCRCNPVIERKNGCMHMKCPQPCLEKMCWNG
CEWNRACMG"

ORIGIN
Query Match 74.0%; Score 348.6; DB 10; Length 1466;
Best Local Similarity 84.2%; Pred. No. 5.3e-103;
Matches 333; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Db 2 GAAGTCCAGAGATGATCAATCTTCAACAGCTTTTATGTATGTGAAAGGCCCTGTC 61
Qy 494 GAGGTCCAGAGCTTAAACCTCACTACACAGCTTTTGTCTACTGCAAGGCCCTGTC 553
Db 62 AAAGATGACAGCCGCGAAACTCAGGTTACAGTCAAGTCAAGGCAAGGCAAGCTCA 121
Qy 554 ACAAGTTCAGCTCGGAACTCCGAGTTCAAGTGGGCACTGCAAGCAAGCAACCTCA 613
Db 122 CCTTGACCCAGGTCATCTTGTCTGGATGATGTTTAAATTCACAAACCGATGATGTG 181
Qy 614 CTTTGCGCCAGGCGCCATCTTGTCTGGATGATGTTTAAATTCACAAACCGATGATGTG 673

| | | | | |
|-----------------------|--|-------------|--|------------------------|
| Db | | 466 | GAGGTCCAGAAAGCTAACCACCACTTACCAAGCTTTTGTGTCATCTGCAGAACGGCCCTGCC | 525 |
| Oy | | 62 | AAGAATGTCAGCCCGGAAAAACTCAGGGTACAGTGACAGCACTCGAGGACGACA | 121 |
| Db | | 526 | ACAAGGTCCAGCCCGGAAAATTCCGAGTTTAGTGCGGCACTCGACACAGCAACCTTCA | 585 |
| Oy | | 122 | CCTTGACCCAGGGGTCATCTTGCTGGGATGATGTTTTTAATTCCAAACCGAGTAGTGATG | 181 |
| Db | | 586 | CTTTGGCCCAAGGCCCCCATCTTGCTGGGATGATGTTCTTAATTCCAAACCGAGTAGTGATG | 645 |
| Oy | | 182 | AATGCCAATCCCACACATGCGCTTGGGACTAGTGACAGAAATTTTCTTTAAATGTGAGCAC | 241 |
| Db | | 646 | AGTSCCAATCTCCAGATCGCCCTGGGACAAAGCTGAATTTTCTTTAAATGTGAGAC | 705 |
| Oy | | 242 | ACCCCACTCTGACAGAAAACATCACTAGCTTTTGCACTGATGCGAACAAATAGTGGGA | 301 |
| Db | | 706 | ACCCAACCTCAGACAAAGACACAATCACTAGCTTTTGAACCTGATCACAACACGCCGA | 765 |
| Oy | | 302 | ACATCACTTGATTAAGTCGACAGACGTCAGAGACCCCGTCCTGATTTTCCAGTGCACT | 361 |
| Db | | 766 | GCATCCCTCGATCGGTGACCGATCGGATCGAAGACCTGTCTTGCTTCCAAATGTAAC | 825 |
| Oy | | 362 | CCGCGCAAGTATTTGCTTAGACTGTTTCCACTTATACTGTGACAAAGACTCAATGATC | 421 |
| Db | | 826 | ACGCGCAAGTATGTTTGGAATGCTGTCCACTGTACTGTGTCACAAGGCTCAACATC | 885 |
| Oy | | 422 | GGCAGTTTGTTCAGACAGCACTTGGGTCATCTCCGCTTGCTGTG | 468 |
| Db | | 886 | GGCAGTTTGTTCAGACAGCACTCAGCTTGGCTTACTGCTGCGGTGTG | 932 |
| RESULT_15 | | | | |
| LOCUS | AF343574 | 1157 bp | mRNA | linear ROD 02-FEB-2002 |
| DEFINITION | Rattus norvegicus parkin transcript variant 6 mRNA, complete cds. | | | |
| ACCESSION | AF343574 | | | |
| VERSION | AF343574.1 | GI:18478865 | | |
| KEYWORDS | . | | | |
| SOURCE | Rattus norvegicus (Norway rat) | | | |
| ORGANISM | Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. | | | |
| REFERENCE | 1 (bases 1 to 1157) D'Agata,V.M., Scapagnini,G. and Cavallaro,S. Functional and molecular diversity of parkin in the rat brain Unpublished | | | |
| AUTHORS | 2 (bases 1 to 1157) D'Agata,V.M., Scapagnini,G. and Cavallaro,S. Direct Submission Submitted (29-JUN-2001) Institute of Biomedging and Pathophysiology of the Central Nervous System, Italian National Research Council (CNR), Piazza Roma, 2, Cacciana 95123, Italy | | | |
| JOURNAL | Location/Qualifiers | | | |
| FEATURES | source | | | |
| CDs | 1..1157 /organism="Rattus norvegicus" /mol_type="mRNA" /db_xref="taxon:10116" 333..1157 /codon_start=1 /product="parkin transcript variant 6" /protein_id="AAU73348.1" /db_xref="GI:18478865" translation="MSGECOSPOCGRAEFPPFGCAHPTSDKDTVALNLITNNSRS IPCLACTDVANNPVLVPCGNRHRYICLDCPLICYTRILNDPOAYVDALGYSLPVAGC PENSILKLHFRIILGEQYINKYIQYGAEBCVLQMGIVLCFRPGCGALFLBQCKRTT CGNGTNGCGFVCRCDEKAHYHEGECDSMEBASGATSQAARVDPRAEAORWESASKE TIKKTYPCRCRWPILEKNGCMHKMCPQPCQKLEMCNMCGCEWRNACMGDHWDV" | | | |
| ORIGIN | | | | |
| Query Match | 73.9%; Score 348.2; DB 10; Length 1157; | | | |
| Best Local Similarity | 84.3%; Pred. No. 7e-103; | | | |

Matches 392; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

| | | | |
|----|-----|---|-----|
| QY | 4 | AGTCACAGAGATGATCAATCTTACACAGCTTTTATGATGATTCACAAGGCCCTGTCAA | 63 |
| Db | 165 | AGTACAGACAGCTAACCCACCTACACAGCTTTTGTCTACTGCAAAAGGCCCTGTCCAC | 224 |
| QY | 64 | AGAGTGCAGCCGGGAAAACCTCAGGGTAACGTGCAGACCTTGACGACAGCAACGCTCAC | 123 |
| Db | 225 | AAGGTCCAGCTGGGAAACCTCCGAGTTCAAGTCGGCACCTGCAGACAAAGCAACCTCAC | 284 |
| QY | 124 | TTGACCCAGGGTCCATCTTGCTGGGATGATGTTTAATTCMAACCGGATGAGTGTGAA | 183 |
| Db | 285 | TTGGCCAGGACCCTATCTTGCTGGGATGATGTCTTAATCCMAACCGGATGAGTGTGAG | 344 |
| QY | 184 | TGCCAATCCCAACACTGCCCCCTGGGACTAGTGCAGAAATTTTCTTAAATGTGAGCAC | 243 |
| Db | 345 | TGCCAATCTTCCAGACTGCCCCCTGGGACAAAGACTGAATTTTCTTAAATGTGAGCAC | 404 |
| QY | 244 | CCCACTCTGACAGAAAGAACATCAGTAGCTTTGCACCTGATCGCACAAAATAGTGGAA | 303 |
| Db | 405 | CCAACCTCAGACAAAGACACATCAGTAGCTTTGAACCTGATCACCAAAACAGCCGAGC | 464 |
| QY | 304 | ATCACTTGCAATTAGTGCACAGAGTCAAGAGCCCCGCTGGTTTCCAGTGAATCC | 363 |
| Db | 465 | ATCCCCCTGCATCGGTGCACGGATGTCAGAAACCTGTGTGCTTCCAAATGTAACAC | 524 |
| QY | 364 | CGCAAGTGAATTTGCTTAGACTGTTCOACTTATATGTTGACAGAGTCAATGATCGG | 423 |
| Db | 525 | CGCCACGTGATCTGTGGATGCTTCCACTTGTACTGTGTACAGAGCTCAAGGATCGG | 584 |
| QY | 424 | CAGTTTGTACGACCCCTCAACTTGCTACTCCCTGCTGTGTG | 468 |
| Db | 585 | CAGTTTGTCCAGACGCTCAGCTTGCTACTCGCTGCCGTGTGTG | 629 |

Search completed: February 26, 2005, 02:12:59
Job time : 2105.69 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2005, 15:46:49 ; Search time 265.776 Seconds
(without alignments)
10490.795 Million cell updates/sec

Title: US-10-622-817-10

Perfect score: 471

Sequence: 1 ggaagtcacagcagtagatc.....accctcgtctgtcgtacg 471

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 471 | 100.0 | 471 | 5 | AAH77664 |
| 2 | 468 | 99.4 | 2960 | 3 | AAA46765 |
| 3 | 466.4 | 99.0 | 2960 | 2 | AAAX99923 |
| 4 | 466.4 | 99.0 | 2960 | 8 | AAD47679 |
| 5 | 466.4 | 99.0 | 2960 | 8 | AAD47680 |
| 6 | 466.4 | 99.0 | 2960 | 8 | ACD13332 |
| 7 | 340.2 | 72.2 | 3092 | 4 | AAFS5253 |
| 8 | 340.2 | 72.2 | 3092 | 4 | AAFS5257 |
| 9 | 340.2 | 72.2 | 3253 | 4 | AAFS5256 |
| 10 | 340.2 | 72.2 | 3253 | 4 | AAFS5254 |
| 11 | 340.2 | 72.2 | 3254 | 4 | AAFS5255 |
| 12 | 340.2 | 72.2 | 3255 | 4 | AAFS5259 |
| 13 | 340.2 | 72.2 | 3255 | 4 | AAFS5244 |
| 14 | 340.2 | 72.2 | 3255 | 4 | AAFS5260 |
| 15 | 340.2 | 72.2 | 3255 | 4 | AAFS5258 |
| 16 | 338.6 | 71.9 | 3014 | 2 | AAFS5247 |
| 17 | 288.4 | 61.2 | 2876 | 2 | AAAX99924 |
| 18 | 288.4 | 61.2 | 2876 | 8 | ACD13333 |
| 19 | 265.4 | 56.3 | 3136 | 4 | AAFS5250 |
| 20 | 262.8 | 55.8 | 2895 | 4 | AAFS5248 |

| | | | | | | |
|----|-------|------|------|----|-----------|---------------------|
| 21 | 251.8 | 53.5 | 2513 | 8 | ACD13334 | ACD13334 Human DNA |
| 22 | 240.8 | 51.1 | 857 | 4 | AAFS5246 | AAFS5246 Nucleotid |
| 23 | 239.8 | 50.9 | 1459 | 4 | AAFS5245 | AAFS5245 Nucleotid |
| 24 | 198.8 | 42.2 | 3170 | 4 | AAFS5251 | AAFS5251 Nucleotid |
| 25 | 100.4 | 21.3 | 318 | 12 | ADNM6486 | Adm96486 Modified |
| 26 | 90.4 | 19.2 | 321 | 12 | ADNM6488 | Adm96488 Modified |
| 27 | 78.8 | 16.7 | 2918 | 4 | AAFS5252 | AAFS5252 Nucleotid |
| 28 | 74 | 15.7 | 518 | 12 | ADNM6490 | Adm96490 Murine pa |
| 29 | 71 | 15.1 | 1458 | 4 | ABL16595 | Ab116595 Drosophi |
| 30 | 53 | 11.3 | 470 | 12 | ADNM6491 | Adm96491 Murine pa |
| 31 | 50.2 | 10.7 | 65 | 6 | ABN53295 | Abn53295 Mouse api |
| 32 | 46.4 | 9.9 | 3865 | 4 | ABL16594 | Ab116594 Drosophi |
| 33 | 35.6 | 7.6 | 1449 | 10 | ADBE07643 | Adbe07643 Novel cod |
| 34 | 34.8 | 7.4 | 361 | 3 | ABAC65989 | Abac65989 Human lun |
| 35 | 34.8 | 7.4 | 361 | 6 | ABLA49208 | Ab149208 Human lun |
| 36 | 34.8 | 7.4 | 361 | 6 | ABO92394 | Abog92394 Human lun |
| 37 | 34.8 | 7.4 | 361 | 9 | ADA28383 | Ada28383 Human lun |
| 38 | 34.8 | 7.4 | 361 | 10 | ADH36947 | Adh36947 Human lun |
| 39 | 34.8 | 7.4 | 405 | 3 | AAAC07249 | Aac07249 Human sec |
| 40 | 34.8 | 7.4 | 405 | 3 | AAAC07249 | Aac07249 Human sec |
| 41 | 33.6 | 7.1 | 362 | 10 | ADFE0743 | Adfe0743 Leukaemia |
| 42 | 33.6 | 7.1 | 395 | 6 | ABO56883 | Abog56883 Human col |
| 43 | 33.6 | 7.1 | 446 | 12 | ADDA0914 | Ada0914 Human CDN |
| 44 | 33.6 | 7.1 | 447 | 9 | ACH26909 | Ach26909 Human adu |
| 45 | 33.6 | 7.1 | 483 | 9 | ACH26764 | Ach26764 Human adu |

ALIGNMENTS

RESULT 1
AAH77664
ID AAH77664 standard; DNA; 471 BP.
XX
XX AAH77664;
XX
XX 13-NOV-2001 (first entry)
XX
XX Nucleotide sequence of human Parkin protein fragment.
XX
XX Human; Parkin-Associated Protein 1; PAPI; Parkin gene;
XX
XX neurodegenerative disease; Parkinson's disease; ss.
XX
XX Homo sapiens.
XX
XX W0200160857-A2.
XX
XX 23-AUG-2001.
XX
XX 15-FEB-2001; 2001WO-FR00461.
XX
XX 17-FEB-2000; 2000FR-00001980.
XX
XX 18-APR-2000; 2000US-0198489P.
XX
XX (AVET) AVENTIS PHARMA SA.
XX
XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
XX Koutnikova H, Brice A, Fournier A, Pradier L, Prades C;
XX
XX Arnould-Reguligne I, Rosier-Montus M, Corti O;
XX
XX WPI; 2001-550047/61.
XX
XX P-PSDB; AAG67213.
XX
XX A new protein, designated Parkin-Associated Protein 1 (PAPI), is an
XX
XX interactor partner of Parkin and is useful to treat neurodegenerative
XX
XX pathologies including Parkinson's disease.
XX
XX Example 1; Page 60-61; 82pp; French.
XX
XX The present sequence encodes the central region of a human Parkin
XX
XX protein. The protein was used to identify Parkin-Associated Protein 1
XX
XX (PAPI) protein. PAPI is associated with the Parkin gene, which is mutated
XX
XX in certain forms of familial (juvenile autosomal recessive) Parkinson's

CC disease. PAPI has some homology with synaptotagmins. PAPI is used to
 CC treat neurodegenerative diseases, particularly to diagnose and treat
 CC Parkinson's disease
 XX

SQ Sequence 471 BP; 118 A; 127 C; 109 G; 117 T; 0 U; 0 Other;

Query Match 100.0%; Score 471; DB 5; Length 471;
 Best Local Similarity 100.0%; Pred. No. 7.8e-153;
 Matches 471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAAGTCCAGCAGGTAGATCAATCTACACACCTTTATGTATGCAAAAGCCCTGT 60
 DB 1 GGAAGTCCAGCAGGTAGATCAATCTACACACCTTTATGTATGCAAAAGCCCTGT 60
 QY 61 CAAAGAGTGCAGCCGGGAAAACCTCAGGGTACAGTGCAGACCTGCAGGCGCAACGCTC 120
 DB 61 CAAAGAGTGCAGCCGGGAAAACCTCAGGGTACAGTGCAGACCTGCAGGCGCAACGCTC 120
 QY 121 ACCTTGACCCAGGGTTCATCTTGGGATGATGTTTAAATTCAAACCGGATGATGAT 180
 DB 121 ACCTTGACCCAGGGTTCATCTTGGGATGATGTTTAAATTCAAACCGGATGATGAT 180
 QY 181 GAATGCCAATCCCACTGACCTGGGACCTAGTGAGAAATTTCTTTAAATGTGAGCA 240
 DB 181 GAATGCCAATCCCACTGACCTGGGACCTAGTGAGAAATTTCTTTAAATGTGAGCA 240
 QY 241 CACCCACCTCTGACAGAAACATCAGTAGCTTGCACCTGATGCACAAATAGTCGG 300
 DB 241 CACCCACCTCTGACAGAAACATCAGTAGCTTGCACCTGATGCACAAATAGTCGG 300
 QY 301 AACATCATTGATTAAGTGCAGACAGCTCAGAGCCCGTCTGTTTCCAGTGCAC 360
 DB 301 AACATCATTGATTAAGTGCAGACAGCTCAGAGCCCGTCTGTTTCCAGTGCAC 360
 QY 361 TCCCGCAGTGAATTTGCTTAGACTGTTTCCATTACTGTGTGACAAAGCTCAATGAT 420
 DB 361 TCCCGCAGTGAATTTGCTTAGACTGTTTCCATTACTGTGTGACAAAGCTCAATGAT 420
 QY 421 CGGCAAGTTTGTTCAGACACCTTCAACTTGGCTACTCCCTGCTGTGTGATG 471
 DB 421 CGGCAAGTTTGTTCAGACACCTTCAACTTGGCTACTCCCTGCTGTGTGATG 471

RESULT 2

AAA46765
 ID AAA46765 standard; cDNA; 2960 BP.

AAA46765;

25-SEP-2000 (first entry)

DE cDNA sequence of the human parkin protein gene.

XX Human; parkin protein; parkin gene; Parkinson's disease;
 KW anti-Parkinson agent; ss.

OS Homo sapiens.

PN W0200031253-A2.

PD 02-JUN-2000.

PF 18-NOV-1999; 99WO-FR002833.

PR 19-NOV-1998; 98FR-00014524.

PR 12-MAR-1999; 99UD-0124239P.

PR 04-AUG-1999; 99FR-00010140.

PA (RHON) RHONE-POULENC ROBER SA.

PI (INRM) INST NAT SANTE & RECH MEDICALE.

XX Brice A, Lucking C, Abbas NE, Denefle P, Ricard S, Bouley S;

DR WPI, 2000-411952/35.

XX New variant forms of the human parkin gene, used as source of primers and
 PT probes for detecting susceptibility to Parkinson's disease.

XX Claim 24; Fig 1; 71pp; French.

PS The present sequence represents a cDNA sequence of the human parkin
 CC protein gene. The specification describes a parkin gene which has genetic
 CC alterations. Cells, or transgenic animals, that express the altered
 CC parkin gene are used to screen for compounds that can counter the effects
 CC of a genetic alteration in the parkin gene, or more generally for
 CC studying the properties of the parkin protein. Detection of the specified
 CC alterations is used to diagnose susceptibility to Parkinson's disease.
 CC The modified polynucleotide is also used to express the corresponding
 CC protein, which is then used to screen for potential anti-Parkinson agents
 CC and to raise antibodies (for detecting variants of parkin protein)

SQ Sequence 2960 BP; 815 A; 726 C; 698 G; 721 T; 0 U; 0 Other;

Query Match 99.4%; Score 468; DB 3; Length 2960;
 Best Local Similarity 100.0%; Pred. No. 2.3e-151;
 Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTCCAGCAGGTAGATCAATCTACACAGCTTTATGTATGCAAAAGCCCTGT 60
 DB 504 GGAAGTCCAGCAGGTAGATCAATCTACACAGCTTTATGTATGCAAAAGCCCTGT 563
 QY 61 CAAAGAGTGCAGCCGGGAAAACCTCAGGGTACAGTGCAGACCTGCAGGCGCAACGCTC 120
 DB 564 CAAAGAGTGCAGCCGGGAAAACCTCAGGGTACAGTGCAGACCTGCAGGCGCAACGCTC 623
 QY 121 ACCTTGACCCAGGGTTCATCTTGGGATGATGTTTAAATTCAAACCGGATGATGAT 180
 DB 121 ACCTTGACCCAGGGTTCATCTTGGGATGATGTTTAAATTCAAACCGGATGATGAT 180
 QY 181 GAATGCCAATCCCACTGACCTGGGACCTAGTGAGAAATTTCTTTAAATGTGAGCA 240
 DB 181 GAATGCCAATCCCACTGACCTGGGACCTAGTGAGAAATTTCTTTAAATGTGAGCA 240
 QY 241 CACCCACCTCTGACAGAAACATCAGTAGCTTGCACCTGATGCACAAATAGTCGG 300
 DB 241 CACCCACCTCTGACAGAAACATCAGTAGCTTGCACCTGATGCACAAATAGTCGG 300
 QY 301 AACATCATTGATTAAGTGCAGACAGCTCAGAGCCCGTCTGTTTCCAGTGCAC 360
 DB 301 AACATCATTGATTAAGTGCAGACAGCTCAGAGCCCGTCTGTTTCCAGTGCAC 360
 QY 361 TCCCGCAGTGAATTTGCTTAGACTGTTTCCATTACTGTGTGACAAAGCTCAATGAT 420
 DB 361 TCCCGCAGTGAATTTGCTTAGACTGTTTCCATTACTGTGTGACAAAGCTCAATGAT 420
 QY 421 CGGCAAGTTTGTTCAGACACCTTCAACTTGGCTACTCCCTGCTGTGTGATG 468
 DB 924 CGGCAAGTTTGTTCAGACACCTTCAACTTGGCTACTCCCTGCTGTGTG 971

RESULT 3

AAK99923
 ID AAK99923 standard; DNA; 2960 BP.

AAK99923;

21-OCT-1999 (first entry)

DE Human parkin gene variant.

XX Parkinson's disease related gene; parkin gene; variant; gene therapy; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
 FH CDS 102..1499


```
PT XX /*tag= a
XX PN MO9940191-A1.
XX 12-AUG-1999.
XX PD
XX PF 09-FEB-1999; 99MO-JP000545.
XX PR 09-FEB-1998; 98JP-00027531.
XX PA (SHIM/) SHIMIZU N.
XX PA (MIZU/) MIZUNO Y.
XX PI Shmizu N, Mizuno Y;
XX DR WPI; 1999-494295/41.
XX DR P-PSDB; AAY32501.
XX PT Gene implicated in the pathology of Parkinson's disease, used for
XX treatment of the disease.
XX PS Claim 1; Page 83-88; 114p; English.
XX CC This sequence represents a gene of the invention, and is implicated in
XX the pathology of Parkinson's disease. This sequence is a variant of the
XX parkin gene found in Parkinson's disease patients. The sequences may be
XX used for the diagnosis, treatment (including gene therapy) and
XX investigation of Parkinson's disease
XX SQ Sequence 2960 BP; 815 A; 727 C; 698 G; 720 T; 0 U; 0 Other;

Query Match 99.0%; Score 466.4; DB 2; Length 2960;
Best Local Similarity 99.8%; Pred. No. 8.4e-151;
Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGTCACAGGTAATCAATCTACACAGCTTTTATGTAATGTAATGCAAGGCCCTGT 60
DB 504 GGAAGTCACAGGTAATCAATCTACACAGCTTTTATGTAATGTAATGCAAGGCCCTGT 563
QY 61 CAAGAAGTCACGCCGGAAAACTCAGGGTACAGTCGACGACCTTCAGAGCGCAACGCTC 120
DB 564 CAAGAAGTCACGCCGGAAAACTCAGGGTACAGTCGACGACCTTCAGAGCGCAACGCTC 623
QY 121 ACCCTGACCCAGGGTCCATCTTGCTGGGATGATGTTTAAATCCAAACCGGATGATGCT 180
DB 624 ACCCTGACCCAGGGTCCATCTTGCTGGGATGATGTTTAAATCCAAACCGGATGATGCT 683
QY 181 GAATGCCAATCCCAACATGAGTGGGACTAGTCAGAAATTTTCTTAAATGTGAGCA 240
DB 684 GAATGCCAATCCCAACATGAGTGGGACTAGTCAGAAATTTTCTTAAATGTGAGCA 743
QY 241 CACCCACCTCTGACAGAAACATCAGTAGCTTTCGACCTGATGCAACAAATAGTCGG 300
DB 744 CACCCACCTCTGACAGAAACATCAGTAGCTTTCGACCTGATGCAACAAATAGTCGG 803
QY 301 AACATCCTTGCAATTAAGTCGACAGAGTCAGAGGCCCTCCGTTTCCATGCAAC 360
DB 804 AACATCCTTGCAATTAAGTCGACAGAGTCAGAGGCCCTCCGTTTCCATGCAAC 863
QY 361 TCCGCGACAGTATTTGCTTGAATGATGTTTCCATTAATAGTGTGACAGCAATGAT 420
DB 864 TCCGCGACAGTATTTGCTTGAATGATGTTTCCATTAATAGTGTGACAGCAATGAT 923
QY 421 CGGAGTTTGTTCAGACCTCAACTTGGCTACTCCCTGCTGTGTG 468
DB 924 CGGAGTTTGTTCAGACCTCAACTTGGCTACTCCCTGCTGTGTG 971

RESULT 4
AAD47679 standard; cDNA; 2960 BP.
XX
AC AAD47679;
```

```
XX XX 24-FEB-2003 (first entry)
XX DT Human Parkin protein encoding cDNA.
XX DE
XX KW Human; Parkin protein; neurological disorder; apoptosis; gene therapy;
XX KW ischaemic stroke; Parkinson's disease; Alzheimer's disease; noctropic;
XX KW transgenic; cerebroprotective; neuroprotective; neurotransplantation;
XX KW gene; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 102..1499
XX FT /*tag= a
XX FT /product= "Human Parkin protein"
XX FT
XX PN MO200279459-A2.
XX PD 10-OCT-2002.
XX PP 02-APR-2002; 2002MO-DK000221.
XX PR 29-MAR-2001; 2001DK-00000525.
XX PR 03-APR-2001; 2001US-0281286P.
XX PA (NSGE-) NSGENE AS.
XX PA Jensen PH;
XX PI
XX DR WPI; 2003-046812/04.
XX DR P-PSDB; AAE30800.
XX XX
XX PT New isolated nucleic acid sequence encoding a Parkin polypeptide, useful
XX for treating, preventing or diagnosing neurological disorders, e.g.,
XX Parkinson's disease, Alzheimer's disease or ischemic stroke, and in
XX screening assays.
XX PT
XX PS Example 1; Page 68-69; 71pp; English.
XX XX
XX CC The invention relates to Parkin protein and its corresponding nucleic
XX acid sequence. The nucleic acid sequence is useful for altering the
XX proteolytic processing of Parkin at its potential cleavage site at Asp
XX 126. The invention is used in manufacturing or testing a pharmaceutical
XX composition for treating and/or preventing a neurological disorder, e.g.,
XX Alzheimer's disease or ischaemic stroke. It also used for detecting the
XX occurrence of proteolytic processing of Parkin at Asp 126 in a sample, in
XX monitoring a potential disposition for a neurodegenerative disease, and
XX for treating, preventing and/or diagnosing Parkinson's disease or other
XX neurodegenerative disorders. The viral vector is used for transforming
XX neuronal cells in vivo or ex vivo. The invention is useful for
XX neurotransplantation into the CNS of a mammal. It may be used in
XX screening assays to identify compounds that increase or decrease
XX apoptosis. It is also used in gene therapy. The present sequence is human
XX Parkin protein encoding cDNA
XX CC
XX SQ Sequence 2960 BP; 815 A; 727 C; 698 G; 720 T; 0 U; 0 Other;

Query Match 99.0%; Score 466.4; DB 8; Length 2960;
Best Local Similarity 99.8%; Pred. No. 8.4e-151;
Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGTCACAGGTAATCAATCTTACACAGCTTTTATGTAATGTAATGCAAGGCCCTGT 60
DB 504 GGAAGTCACAGGTAATCAATCTTACACAGCTTTTATGTAATGTAATGCAAGGCCCTGT 563
QY 61 CAAGAAGTCACGCCGGAAAACTCAGGGTACAGTCGACGACCTTCAGAGCGCAACGCTC 120
DB 564 CAAGAAGTCACGCCGGAAAACTCAGGGTACAGTCGACGACCTTCAGAGCGCAACGCTC 623
QY 121 ACCCTGACCCAGGGTCCATCTTGCTGGGATGATGTTTAAATCCAAACCGGATGATGCT 180
DB 624 ACCCTGACCCAGGGTCCATCTTGCTGGGATGATGTTTAAATCCAAACCGGATGATGCT 683
```

QY 181 GAATGCCAATCCCACTGCTCCCTGGGACTAGTGCAGAAATTTTCTTAAATGTGAGCA 240
DB 684 GAATGCCAATCCCACTGCTCCCTGGGACTAGTGCAGAAATTTTCTTAAATGTGAGCA 743
QY 241 CACCCCACTCTGACAGAGAAACATCAGTAGCTTTTGACCTGATCGCACAAAATAGTCGG 300
DB 744 CACCCCACTCTGACAGAGAAACATCAGTAGCTTTTGACCTGATCGCACAAAATAGTCGG 803
QY 301 AACATCACTTGACTTACGTCAGACGTCAGAGAGCCCGCTGCTGTTTCCAGTGCAC 360
DB 804 AACATCACTTGACTTACGTCAGACGTCAGAGAGCCCGCTGCTGTTTCCAGTGCAC 863
QY 361 TCCCGCAGCTGATTTGCTTAACTGTTTCACTTATAGTGTGACAAAGACTCAATGAT 420
DB 864 TCCCGCAGCTGATTTGCTTAACTGTTTCACTTATAGTGTGACAAAGACTCAATGAT 923
QY 421 CGGCACTTTGTTCAAGACCTCTCACTGCTACTCCCTGCTTGTGTG 468
DB 924 CGGCACTTTGTTCAAGACCTCTCACTGCTACTCCCTGCTTGTGTG 971

RESULT 5

ID AAD47680 standard; DNA; 2960 BP.

AC AAD47680;

DT 24-FEB-2003 (first entry)

DE Human Parkin D126 mutant DNA.

KM Human; Parkin protein; neurological disorder; apoptosis; gene therapy;
KM ischaemic stroke; Parkinson's disease; Alzheimer's disease; nocturnal;
KM transgenic; cerebroprotective; neuroprotective; neurotransplantation;
KM mutant; gene; ds.

OS Homo sapiens.
OS Synthetic.

XX Key Location/Qualifiers

FT CDS 102..1499

FT /product= "Human Parkin D126 mutant protein"

FT /trans_except= (pos:768..770, aa:Ser)

XX MO200279459-A2.

PD 10-OCT-2002.

PF 02-APR-2002; 2002MO-DK000221.

XX 29-MAR-2001; 2001DK-0000525.

PR 03-APR-2001; 2001US-0281286P.

XX (NSGE-) NSGENE AS.

XX Jensen PH;

XX WPI; 2003-046812/04.

XX P-PSDB; AAE30801.

XX New isolated nucleic acid sequence encoding a Parkin polypeptide, useful
XX for treating, preventing or diagnosing neurological disorders, e.g.
XX Parkinson's disease, Alzheimer's disease or ischemic stroke, and in
XX screening assays.

XX Claim 4; Page 69-71; 71pp; English.

XX The invention relates to Parkin protein and its corresponding nucleic
XX acid sequence. The nucleic acid sequence is useful for altering the
XX proteolytic processing of Parkin at its potential cleavage site at Asp
XX 126. The invention is used in manufacturing or testing a pharmaceutical

CC composition for treating and/or preventing a neurological disorder, e.g.
CC Alzheimer's disease or ischaemic stroke. It also used for detecting the
CC occurrence of proteolytic processing of Parkin at Asp 126 in a sample, in
CC monitoring a potential disposition for a neurodegenerative disease, and
CC for treating, preventing and/or diagnosing Parkinson's disease or other
CC neurodegenerative disorders. The viral vector is used for transforming
CC neuronal cells in vivo or ex vivo. The invention is useful for
CC neurotransplantation into the CNS of a mammal. It may be used in
CC screening assays to identify compounds that increase or decrease
CC apoptosis. It is also used in gene therapy. The present sequence is human
CC Parkin D126 mutant DNA

Seq Sequence 2960 BP; 816 A; 726 C; 698 G; 720 T; 0 U; 0 Other;

Query Match 99.0%; Score 466.4; DB 8; Length 2960;

Best Local Similarity 99.8%; Pred. No. 8,4e-151;

Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGTCACAGAGTATGATCAATCTACACAGCTTTTATGTATGTCAAAGGCCCTGT 60
DB 504 GGAAGTCACAGAGTATGATCAATCTACACAGCTTTTATGTATGTCAAAGGCCCTGT 563
QY 61 CAAAGAGTGCAGCCGGGAAAACACAGGTAAGTGCAGACCTGACAGGCAAGCAAGCTC 120
DB 564 CAAAGAGTGCAGCCGGGAAAACACAGGTAAGTGCAGACCTGACAGGCAAGCAAGCTC 623
QY 121 ACCTTGACCCAGGCTCATCTTGTGAGTATGTTTAAATTCAAACCGATAGTGT 180
DB 624 ACCTTGACCCAGGCTCATCTTGTGAGTATGTTTAAATTCAAACCGATAGTGT 683
QY 181 GAATGCCAATCCCACTGCTCCCTGGGACTAGTGCAGAAATTTTCTTAAATGTGAGCA 240
DB 684 GAATGCCAATCCCACTGCTCCCTGGGACTAGTGCAGAAATTTTCTTAAATGTGAGCA 743
QY 241 CACCCCACTCTGACAGAGAAACATCAGTAGCTTTTGACCTGATCGCACAAAATAGTCGG 300
DB 744 CACCCCACTCTGACAGAGAAACATCAGTAGCTTTTGACCTGATCGCACAAAATAGTCGG 803
QY 301 AACATCACTTGACTTACGTCAGACGTCAGAGAGCCCGCTGCTGTTTCCAGTGCAC 360
DB 804 AACATCACTTGACTTACGTCAGACGTCAGAGAGCCCGCTGCTGTTTCCAGTGCAC 863
QY 361 TCCCGCAGCTGATTTGCTTAACTGTTTCACTTATAGTGTGACAAAGACTCAATGAT 420
DB 864 TCCCGCAGCTGATTTGCTTAACTGTTTCACTTATAGTGTGACAAAGACTCAATGAT 923
QY 421 CGGCACTTTGTTCAAGACCTCTCACTGCTACTCCCTGCTTGTGTG 468
DB 924 CGGCACTTTGTTCAAGACCTCTCACTGCTACTCCCTGCTTGTGTG 971

RESULT 6

ID ACD13332 standard; cDNA; 2960 BP.

AC ACD13332;

DT 13-AUG-2003 (first entry)

DE Human DNA encoding a p53 modifier. SEQ ID 1.

KM Human; ss; gene; p53 modifier; cytosolic; cancer; cytosolic;
KM antiangiogenic; antiapoptotic; p53 pathway; breast cancer; colon cancer;
KM kidney cancer; lung cancer; ovarian cancer; angiogenesis; cell cycle;
KM apoptotic disorder; cell proliferation disorder.

XX Homo sapiens.

XX WO200299122-A1.

XX 12-DEC-2002.

XX 03-JUN-2002; 2002MO-US017382.

XX 05-JUN-2001; 2001US-0296076P.
 PR 10-OCT-2001; 2001US-0328605P.
 PR 15-FEB-2002; 2002US-0357253P.
 XX (EXEL-) EXELIXIS INC.
 PA
 PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP,
 DR P-PSDB; ABO07156.
 XX WPI; 2003-156859/15.
 PT Identifying modulators of the p53 pathway for use in treating apoptotic
 PT or cell proliferation disorders, comprises screening for agents that
 PT modulate activity of a human ortholog of genes that modify the p53
 PT pathway in Drosophila.
 XX
 PS Example 2; Page 166-167; 678pp; English.
 XX
 CC The invention relates to identifying (M1) a candidate p53 pathway
 CC modulating agent, by contacting an assay system comprising a purified HM
 CC polypeptide (human orthologue of genes that modify the p53 pathway in
 CC Drosophila) or nucleic acid with a test agent under conditions, where but
 CC for the presence of the test agent, the system provides a reference
 CC activity, and detecting a test agent-biased activity of the assay system.
 CC Also included are modulating (M2) a p53 pathway of a cell (comprising
 CC contacting a cell defective in p53 function with a candidate modulator
 CC that specifically binds to a HM polypeptide comprising an HM amino acid
 CC sequence, where p53 function is restored), modulating (M3) a p53 pathway
 CC in a mammalian cell (comprising contacting the cell with an agent that
 CC specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
 CC a disease in a patient (comprising: (a) obtaining a biological sample
 CC from the patient; (b) contacting the sample with a probe for HM
 CC expression; (c) comparing the results with a control; and (d) determining
 CC whether the comparison indicates a likelihood disease). (M1) is useful
 CC for identifying modulators of the p53 pathway. A probe for HM expression
 CC is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,
 CC in a patient, where the cancer has greater than 25 % expression level.
 CC Modulators identified by (M1) are useful in a variety of diagnostic and
 CC therapeutic applications, where disease or disorder prognosis is related
 CC to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
 CC proliferation disorders (e.g. cancer). Another two new methods (M2 and
 CC M3) are useful for modulating the p53 pathway of a cell, thus restoring
 CC the p53 function of the cell, so that the cell undergoes normal
 CC proliferation or progression through the cell cycle. (M2) and (M3) are
 CC also useful for treating defects in the p53 pathway such as angiogenic,
 CC apoptotic or cell proliferation disorders. The present sequence is an HM
 CC nucleic acid encoding a p53 pathway modifying protein
 XX
 S0 Sequence 2960 BP; 815 A; 727 C; 698 G; 720 T; 0 U; 0 Other;
 Query Match 99.0%; Score 466.4; DB 8; Length 2960;
 Best Local Similarity 99.8%; Pred. No. 8.4e-15;
 Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGAAGTCCAGAGGTAATGTAATCTACAACAGCTTTATGTTATGTAAGGCCCTCT 60
 DB 504 GGAAGTCCAGAGGTAATGTAATCTACAACAGCTTTATGTTATGTAAGGCCCTCT 563
 QY 61 CAAAGATGACGCGGAAAACTCAGGGTACAGTGCACACTCAGGCAAGCAACGCTC 120
 DB 564 CAAAGATGACGCGGAAAACTCAGGGTACAGTGCACACTCAGGCAAGCAACGCTC 623
 QY 121 ACCTTGACCCAGGGTCCATCTTGGTGGATGATGTTTAAATCCAAACCGGATGATGCT 180
 DB 624 ACCTTGACCCAGGGTCCATCTTGGTGGATGATGTTTAAATCCAAACCGGATGATGCT 683
 QY 181 GAAAGCCAAATCCCAACATGCGCTGGGACTAGTGAAGAAATTTTCTTAAATGAGGACA 240
 DB 684 GAAAGCCAAATCCCAACATGCGCTGGGACTAGTGAAGAAATTTTCTTAAATGAGGACA 743
 QY 241 CACCCCACTTGACAAGGAACAATCAAGTACCTTGACCTGATCGAACAATAAGTGGG 300

DB 744 CACCCCACTTGACAAGGAACAACACAGTACCTTGACCTGATCGAACAATAAGTGGG 803
 QY 301 AACATGACTTGACATTAAGTGACAGAGCTGAGAGCCCGCTGTTTCCAGTGCAAC 360
 DB 804 AACATGACTTGACATTAAGTGACAGAGCTGAGAGCCCGCTGTTTCCAGTGCAAC 863
 QY 361 TCCCGCACGATGATTTTGCTTAGACTGTTTCCACTTATATCTGTGACAAAGACTCAATGAT 420
 DB 864 TCCCGCACGATGATTTTGCTTAGACTGTTTCCACTTATATCTGTGACAAAGACTCAATGAT 923
 QY 421 CGGCAAGTTTGTTCACGACACCTCAACTTGGCTACTTCCTGCTTGTTG 468
 DB 924 CGGCAAGTTTGTTCACGACACCTCAACTTGGCTACTTCCTGCTTGTTG 971
 RESULT 7
 AAF55253
 ID AAF55253 standard; cDNA; 3043 BP.
 XX
 AC AAF55253;
 XX
 DT 29-MAY-2001 (first entry)
 XX
 XX Nucleotide sequence of murine parkin2 with a frameshift mutation.
 XX
 KW Parkinson's disease; 6q25.2-27; neurodegenerative disease;
 KW Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
 KW Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
 KW brain tumour; head trauma; stroke; vascular irregularity;
 KW metabolic irregularity; ss.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 129..1019
 FT /*tag= a
 FT /product= "truncated parkin2"
 FT
 XX
 XX EPI081225-A1.
 XX
 PD 07-MAR-2001.
 XX
 PF 30-AUG-1999; 99BP-00116766.
 PF
 XX 30-AUG-1999; 99BP-00116766.
 PR
 XX 30-AUG-1999; 99BP-00116766.
 PA (BIOF-) BIOFRONTIERA PHARM GMBH.
 PI
 PI Luebbert H;
 XX
 XX WPI; 2001-212797/22.
 DR P-PSDB; AAB67526.
 DR
 PT New polynucleotides encoding mouse parkin2 protein, useful for producing
 PT a transgenic non-human animal as an animal model for neurodegenerative
 PT diseases.
 PT
 PS Claim 3; Page 29-30; 62pp; English.
 XX
 XX The present sequence encodes a murine parkin2 polypeptide. The
 XX polynucleotide sequence contains a frameshift mutation, leading to a
 XX truncated protein. Mutations or deletions in the parkin2 gene cause
 XX Parkinson's disease in humans. The human parkin2 gene is located in gene
 XX region 6q25.2-27. Parkin2 polypeptides and polynucleotides are useful for
 XX analysing neurodegenerative diseases. They are also useful for testing
 XX the efficacy of the treatment of a neurodegenerative disease such as
 XX Parkinson's disease, Alzheimer's disease, Huntington's disease,
 XX amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease,
 XX Pick's disease, Prion disease, and secondary causes inducing Parkinson's
 XX syndromes like toxins, drugs, brain tumours, head trauma, stroke,
 XX CC vascular irregularities or metabolic irregularities, associated with a
 XX less active or non-active parkin protein

| | | |
|---|---|--|
| SQL | Sequence | 3043 BP; 806 A; 763 C; 737 G; 737 T; 0 U; 0 Other; |
| | Query Match | 72.2%; Score 340.2; DB 4; Length 3043; |
| | Best Local Similarity | 84.0%; Pred. No. 7e-107; |
| | Matches | 384; Conservative 0; Mismatches 73; Indels 0; Gaps 0 |
| OY | 11 | CAGTAGATCATCTTCAACAACAGCTTTTATGTGTAATTGCAAAAGCCCTGTCAAAGATGC 70 |
| Dd | 538 | CAGTTAAACCACCTTCAACACAGCTTTTTTATCTACTGCAAAAGGCCCTTGCCAAAGATCC 597 |
| OY | 71 | AGCCGGGAAAACTCAGGGGTACAGTSCAAGCAGCACTTCAGGCGAACAAOGCTCACCTTGACCC 130 |
| Dd | 598 | AGCCTGGAAAGCTCCGAGTTCAGTGTGGCACCCTGCAAAACGAACCAACCTCACCTTGACCC 657 |
| OY | 131 | AGGGTCACCTTTGGCTGGGATGTATGTTTTAAATTTCCAAACCGGATGTAGTGTAAATGCCAAT 190 |
| Dd | 658 | AGGGCCCATCTGTCTGGGACGATGTCTTAATTTCCAACCGGATGTAGTGTAGTGCAGT 717 |
| OY | 191 | CCCCACACTGCCCTGGGACTATGTGAGAATTTTTCTTTAAATGTGAGACACACCCCACT 250 |
| Dd | 718 | CTCCGAGATCGCCCTGGAAACAAGCTGAATTTTTCTTTAAATGTGAGACACACCCCACT 777 |
| OY | 251 | CTGACAAAGGAAACATCAGTAGCTTTGCACCTGATGCAACAAATATGTGGAAACATCACTT 310 |
| Dd | 778 | CAGCAAAAGAACACGTCGGTAGCTTTGAACTGTATCACCAGCAACAGGGGACGATCCCTT 837 |
| OY | 311 | GCATTTACGTGCACAGACGCTCAGAGAGCCCGCTGTGTTTTTCCAGTGAACCTCCGGCCACG 370 |
| Dd | 838 | GCATGACGTGCACAGATGTCAAGAGCCCTGTGCTGTCTTCCAGTGTAAACAACGCTCACG 897 |
| OY | 371 | TGATTTGCTTAGACTGTTCACATTAACTGTGTGACAAAGATCAATGATGGCAGATTG 430 |
| Dd | 898 | TGATCTGTTTGACAGCTTTCCACTGTATGTGTGTACAAAGACTCAACGATGGGAGTTTG 957 |
| OY | 431 | TTCAAGACCCCTCAACCTTGGCTACTCCCTGGCTTGCT 467 |
| Dd | 958 | TCACGATGCTCAACTTGGCTACTCCCTGGCTTGCT 994 |
| RESULT 8 | | |
| AAPF5257 | ID | AAPF5257 standard; cDNA: 3092 BP. |
| XX | AAAF5257; | |
| AC | | |
| XX | | |
| XX | 29-MAY-2001 | (first entry) |
| DT | | |
| XX | | |
| XX | | |
| Nucleotide sequence of murine parkin2 with a frameshift mutation. | | |
| KW | Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease; | |
| KM | Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis; | |
| KM | Multicystic atrophy; Wilson's disease; Pick's disease; Prion disease; | |
| KW | brain tumour; head trauma; stroke; vascular irregularity; | |
| KM | metabolic irregularity; ss. | |
| XX | | |
| OS | Mus sp. | |
| XX | | |
| FH | Key | Location/Qualifiers |
| FT | CDS | 129..173 |
| FT | | /tag= a |
| FT | | /product= "truncated parkin2" |
| PN | EP1081225-A1. | |
| XX | | |
| PD | 07-MAR-2001. | |
| XX | | |
| PF | 30-AUG-1999; | 99EP-00116766. |
| XX | | |
| PR | 30-AUG-1999; | 99EP-00116766. |
| PA | (BIOF-) BIOFRONTIERA PHARM GMBH. | |
| XX | | |
| Luebbert H; | | |
| II | | |

XX WP1: 2001-212797/22.
DR P-PSDB; AAB67550.
XX
XX New polynucleotides encoding mouse parkin2 protein, useful for producing
PT a transgenic non-human animal as an animal model for neurodegenerative
PT diseases.

XX
XX Claim 3, Page 34-35, 62pp; English.

XX
XX The present sequence encodes a murine parkin2 polypeptide. The
CC polynucleotide sequence contains a frameshift mutation, leading to a
CC truncated protein. Mutations or deletions in the parkin2 gene cause
CC Parkinson's disease in humans. The human parkin2 gene is located in gene
CC region 6q25.2-77. Parkin2 polypeptides and polynucleotides are useful for
CC analysing neurodegenerative diseases. They are also useful for testing
CC the efficacy of the treatment of a neurodegenerative disease such as
CC Parkinson's disease, Alzheimer's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, Multi-system atrophy, Wilson's disease,
CC Pick's disease, Prion disease, and secondary causes inducing Parkinson's
CC syndromes like toxins, drugs, brain tumours, head trauma, stroke,
CC vascular irregularities or metabolic irregularities, associated with a
CC less active or non-active parkin protein

XX
XX Sequence 3092 BP, 818 A, 778 C, 761 G, 735 T, 0 U, 0 Other;

| | | | | | | |
|----------|---|---|-------------------|-----------|--------------|--|
| | Query Match | 72.2% | Score 340.2; | DB 4; | Length 3092; | |
| | Best Local Similarity | 84.0%; | Pred. No. 7e-107; | | | |
| | Matches 384; | Conservative 0; | Mismatches 73; | Indels 0; | Gaps 0 | |
| OY | 11 | CAGGTAGATCAATCTACAAAGCTTTTATGTAATGGTCAAAGGCCCTGTCCAAAGATGC | 70 | | | |
| Db | 375 | CAGTTAAACCACCTTACAACAGCTTTTTTCATCTACTCGAAGAAGGCCCTTGCCAAGAATCC | 434 | | | |
| OY | 71 | AGCCGGGAAAACCTCAGGGGTACAGTGCAGCATCTGAGGCAAGCAACGCTCACCTTGACCC | 130 | | | |
| Db | 435 | AAGCTGGAAAGCTCCGAATTGATGTGGCACCCTGCAACAAAGCAACCTTCACTTTGGCCC | 494 | | | |
| OY | 131 | AGGGTCATCTTGGCTGGAGATGTTTTTAATTCMAAACCGGATGAGTGAATGCCAAT | 190 | | | |
| Db | 495 | AGGGCCCATCTTGGCTGGGACGATGCTTAATTCMAAACGGATGAGTGAATGCCAAT | 554 | | | |
| OY | 191 | CCCCACATGCCCCCTGGGACATAAGTCAGAAATTTTTCTTTAAATGTGGAGACAACCCCACT | 250 | | | |
| Db | 555 | CTCAGACATGCCCCCTGGGAAACCAAGACTGAATTTTTCTTTAAATGTGGAGACAACCCCACT | 614 | | | |
| OY | 251 | CTGACAAGGAACATCATAGTAGCTTTGCACTCGATGCGAAACAATAGTGGAAATCACTT | 310 | | | |
| Db | 615 | CAGACAAAGACACGTCGGTAGCTTTGAACTCGATACACGACAAAGGCGCAGATCCCTT | 674 | | | |
| OY | 311 | GCATTAACGTGCACAGACGTGAGAGGCCCGCTCGTGGTTTTTCCAGTGCAACTCCGCGCAG | 370 | | | |
| Db | 675 | GCATAGGTGTGACAGATGTGAGAGGCCCTGTCCTGATCTTCAAGTGAACACACGTCACG | 734 | | | |
| OY | 371 | TGATTTGCTTAGATATGTTTTCCACTTAATACGTGTGACAAGACTCAATGATGGCAGTTTG | 430 | | | |
| Db | 735 | TGATCTTTTGGACTGTGTTTCCATTGTAATGTGTCAACAAGCTCAACGATGGCAGTTTG | 794 | | | |
| OY | 431 | TTCAAGCAACCTCAACTGGCTACTCCCTGCCTTGTGT | 467 | | | |
| Db | 795 | TCCACGATGCTCAACTGGCTACTCCCTGCCTGTGT | 831 | | | |
| RESULT 9 | | | | | | |
| AAF55256 | | | | | | |
| ID | AAF55256 | standard; cDNA; 3253 BP. | | | | |
| XX | AAF55256; | | | | | |
| XX | AC | | | | | |
| XX | AAF55256; | | | | | |
| DT | 29-MAY-2001 | (first entry) | | | | |
| DE | Nucleotide sequence of murine parkin2 with a frameshift mutation. | | | | | |

KM Parkinson's disease; 6q25.2-27; neurodegenerative disease;
 KM Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
 KM Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
 KM brain tumour; head trauma; stroke; vascular irregularity;
 KM metabolic irregularity; ss.
 XX
 OS Mus sp.
 FH Key Location/Qualifiers
 FT CDS 129..362
 FT /*tag= a
 FT /product= "truncated parkin2"
 PN EPI081225-A1.
 PD 07-MAR-2001.
 XX 30-AUG-1999; 99EP-00116766.
 XX 30-AUG-1999; 99EP-00116766.
 PR (BIOF-) BIOFRONTIERA PHARM GMBH.
 PA Luebbert H;
 PI
 XX
 XX WPI: 2001-212797/22.
 DR P-PSDB; AAB67529.
 XX
 PT New polynucleotides encoding mouse parkin2 protein, useful for producing
 PT a transgenic non-human animal as an animal model for neurodegenerative
 PT diseases.
 PT
 XX
 XX Claim 3; Page 33-34; 62pp; English.
 XX
 XX The present sequence encodes a murine parkin2 polypeptide. The
 CC polynucleotide sequence contains mutations, causing Arg78 to be changed
 CC to a stop codon, leading to a truncated protein. Mutations or deletions
 CC in the parkin2 gene cause Parkinson's disease in humans. The human
 CC parkin2 gene is located in gene region 6q25.2-27. Parkin2 polypeptides
 CC and polynucleotides are useful for analysing neurodegenerative diseases.
 CC They are also useful for testing the efficacy of the treatment of a
 CC neurodegenerative disease such as Parkinson's disease, Alzheimer's
 CC disease, Huntington's disease, amyotrophic lateral sclerosis, Multi-
 CC system atrophy, Wilson's disease, Pick's disease, Prion disease, and
 CC secondary causes inducing Parkinson's syndromes like toxins, drugs, brain
 CC tumours, head trauma, stroke, vascular irregularities or metabolic
 CC irregularities, associated with a less active or non-active parkin
 CC protein.
 CC
 XX
 XX Sequence 3253 BP; 851 A; 818 C; 808 G; 776 T; 0 U; 0 Other;
 S0
 Query Match 72.2%; Score 340.2; DB 4; Length 3253;
 Best Local Similarity 84.0%; Pred. No. 7.2e-107;
 Matches 384; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
 QY 11 CAGGTAGTCAATATACAAAGCTTTTATGTGTATGCAAAAGGCCCTGTAAAGAGTGC 70
 DB 536 CAGTTAAACCCACCTACAAAGCTTTTATCATCTACTGCAAAAGGCCCTGTCAAGGTCC 595
 QY 71 AGCCGGGAAAACTAGAGGTACAGTGCAGACCTGCAGGAGCAAGCTCACTTGAACC 130
 DB 596 AGCCTGAAAGCTCTCGAGATTGATGTGGACCTGCAAAAGCAACCTCCTTGGCCC 655
 QY 131 AGGCTCATCTTGTGTGGAGTATTTTAAATTCGAACCGGATGAGTGAATGCCAT 190
 DB 656 AGGCGCCATCTTGTGTGGAGCATGTCTTAAATTCGAACCGGATGAGTGAATGCCAGT 715
 QY 191 CCCCACTGCTGCTGGAGTACTAGTCAAAATTTTCTTAAATGTGGAGCAACCCCACT 250
 DB 716 CTCAGACTGCCCTGGAAACGAGACTGAATTTTCTTAAATGTGGAGCAACCCCACT 775
 QY 251 CTGACAGAAACATCACTAGCTTTGCACTGATCGCAAAATAGTCGGAACATCACTT 310
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

DB 776 CAGACAAAGACACCTCGTAGCTTTGAAACCTGATCACCAGCAAGCGCAGATCCCTT 835
 QY 311 GCATTAGTGACAGAGCTGAGAGCGCCGTCCTGTTTTCAGTGAACCTCCGCCACG 370
 DB 836 GCATAGCGTCACAGATGTACAGAGCCCTGCTGCTGCTTCCAGTGTAAACCGCTCACG 895
 QY 371 TGAATTCCTTGAAGCTGTTTCCACTTATATCTGTGACAGACTCAATGATGGAGTTTG 430
 DB 896 TGATCTGTTGAGACTGTTTCCACTTGTATGTGTGCACAAAGCTCAAGATGGCAGTTTG 955
 QY 431 TTCACGACCTCAACTTGGCTACTCCCTGCTGTGT 467
 DB 956 TCACGATGCTCAACTTGGCTACTCCCTGCTGTGT 992
 RESULT 10
 AAF55254
 ID AAF55254 standard; cDNA; 3253 BP.
 XX
 XX AAF55254;
 AC
 XX 29-MAY-2001 (first entry)
 DT
 XX
 XX Nucleotide sequence of murine parkin2 with a frameshift mutation.
 DE
 XX
 KM Parkinson's disease; 6q25.2-27; neurodegenerative disease;
 KM Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
 KM Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
 KM brain tumour; head trauma; stroke; vascular irregularity;
 KM metabolic irregularity; ss.
 KM
 XX
 XX Mus sp.
 OS
 XX
 XX Key Location/Qualifiers
 FH CDS 129..242
 FT /*tag= a
 FT /product= "truncated parkin2"
 PT
 PT
 XX
 XX EPI081225-A1.
 XX
 XX 07-MAR-2001.
 PD
 XX 30-AUG-1999; 99EP-00116766.
 PF
 XX 30-AUG-1999; 99EP-00116766.
 PR
 XX (BIOF-) BIOFRONTIERA PHARM GMBH.
 PA Luebbert H;
 PI
 XX
 XX WPI: 2001-212797/22.
 DR P-PSDB; AAB67529.
 XX
 PT New polynucleotides encoding mouse parkin2 protein, useful for producing
 PT a transgenic non-human animal as an animal model for neurodegenerative
 PT diseases.
 PT
 XX
 XX Claim 3; Page 30-31; 62pp; English.
 XX
 XX The present sequence encodes a murine parkin2 polypeptide. The
 CC polynucleotide sequence contains mutations, causing Gln38 to be changed
 CC to a stop codon, leading to a truncated protein. Mutations or deletions
 CC in the parkin2 gene cause Parkinson's disease in humans. The human
 CC parkin2 gene is located in gene region 6q25.2-27. Parkin2 polypeptides
 CC and polynucleotides are useful for analysing neurodegenerative diseases.
 CC They are also useful for testing the efficacy of the treatment of a
 CC neurodegenerative disease such as Parkinson's disease, Alzheimer's
 CC disease, Huntington's disease, amyotrophic lateral sclerosis, Multi-
 CC system atrophy, Wilson's disease, Pick's disease, Prion disease, and
 CC secondary causes inducing Parkinson's syndromes like toxins, drugs, brain
 CC tumours, head trauma, stroke, vascular irregularities or metabolic
 CC irregularities, associated with a less active or non-active parkin
 CC protein.

XX SQ Sequence 3253 BP; 851 A; 818 C; 808 G; 776 T; 0 U; 0 Other;
Query Match 72.2%; Score 340.2; DB 4; Length 3253;
Best Local Similarity 84.0%; Pred. No. 7.2e-107;
Matches 384; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
QY 11 CAGGTAGATCAATCTACAAAGCTTTTATGTTATGTAATGCAAAAGGCCCTGTCAAGAGTGC 70
DB 536 CAGTTAAACCACTACACAGCTTTTTCATCTACTGAAAGGCCCTGTCCACAAAGTCC 595
QY 71 AGCCGGGAAAATCTCAGGAGTACAGTGCAGACCTTGAGGAGGCAACGCTCACTTGACCC 130
DB 596 AGCTGGAAAGCTCCGAGTTCACTGTGGCACCTGCAACCAAGAACCTTCACCTTGAGCC 655
QY 131 AGGGCCATCTTGTGGGATGATGTTTAAATCCAAACCGGATGATGATGATGATGATGAT 190
DB 656 AGGGCCATCTTGTGGGATGATGTTTAAATCCAAACCGGATGATGATGATGATGATGAT 715
QY 191 CCCCACTGCTGGGAGTACAGTGCAGAAATTTTCTTAAATGAGACACCCCACT 250
DB 716 CTCGAGACTGCTGGGAGTACAGTGCAGAAATTTTCTTAAATGAGACACCCCACT 775
QY 251 CTGACAAAGAAACATCAGTACCTTTCAGCTGATGCAACAAATAGTGGAACTCACTT 310
DB 776 CAGACAAAGAACCTCGGTAGCTTGAACCTGATCACCAGCAAGGCGCAGCATCCCTT 835
QY 311 GCATTAGGTGACAGAGTGCAGAGCCCGTCTGTTTTCAGTGCAGACCTCCGCGACG 370
DB 836 GCATTAGGTGACAGAGTGCAGAGCCCGTCTGTTTTCAGTGCAGACCTCCGCGACG 895
QY 371 TGATTTGCTTACAGTGTTCCTTCACTTATGAGTGCAGAACTCAATGATGAGGAGTTTG 430
DB 896 TGATCTGTTTGAAGTGTTCCTTCACTTATGAGTGCAGAACTCAATGATGAGGAGTTTG 955
QY 431 TTCAGACCTCTCAACTTGGCTACTTCCCTGCTGTGT 467
DB 956 TTCAGATGCTCAACTTGGCTACTTCCCTGCTGTGT 992
RESULT 11
AAFS5255
ID AAF55255 standard; cDNA; 3254 BP.
XX AC AAF55255;
XX DT 29-MAY-2001 (first entry)
XX DE Nucleotide sequence of murine parkin2 with a frameshift mutation.
XX KM Parkinson's disease; 6q25.2-27; neurodegenerative disease;
XX KM Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
XX KM Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
XX KM brain tumour; head trauma; stroke; vascular irregularity;
XX KM metabolic irregularity; ss.
XX OS Mus sp.
XX FT Key Location/Qualifiers
XX FT CDS 129..290
XX FT /*tag= a
XX FT /product= "truncated parkin2"
XX PN BP1081225-A1.
XX PD 07-MAR-2001.
XX PF 30-AUG-1999; 99EP-00116766.
XX PR 30-AUG-1999; 99EP-00116766.
XX PA (BIOF-) BIOFRONTIERA PHARM GMBH.
XX

PI Luebert H;
XX WPI: 2001-212797/22.
DR P-PSDB; AAB67528.
XX
PT New polynucleotides encoding mouse parkin2 protein, useful for producing
PT a transgenic non-human animal as an animal model for neurodegenerative
PT diseases.
PS Claim 3; Page 31-32; 62pp; English.
XX
XX The present sequence encodes a murine parkin2 polypeptide. The
CC polynucleotide sequence contains mutations, causing Asn54 to be changed
CC to a stop codon, leading to a truncated protein. Mutations or deletions
CC in the parkin2 gene cause Parkinson's disease in humans. The human
CC parkin2 gene is located in gene region 6q25.2-27. Parkin2 polypeptides
CC and polynucleotides are useful for analysing neurodegenerative diseases.
CC They are also useful for testing the efficacy of the treatment of a
CC neurodegenerative disease such as Parkinson's disease, Alzheimer's
CC disease, Huntington's disease, amyotrophic lateral sclerosis, Multi-
CC system atrophy, Wilson's disease, Pick's disease, Prion disease, and
CC secondary causes inducing Parkinson's syndromes like toxins, drugs, brain
CC tumours, head trauma, stroke, vascular irregularities or metabolic
CC irregularities, associated with a less active or non-active parkin
CC protein
XX
SQ Sequence 3254 BP; 851 A; 818 C; 809 G; 776 T; 0 U; 0 Other;
Query Match 72.2%; Score 340.2; DB 4; Length 3254;
Best Local Similarity 84.0%; Pred. No. 7.2e-107;
Matches 384; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
QY 11 CAGGTAGATCAATCTACAAAGCTTTTATGTTATGTAATGCAAAAGGCCCTGTCAAGAGTGC 70
DB 537 CAGTTAAACCACTACACAGCTTTTTCATCTACTGAAAGGCCCTGTCCACAAAGTCC 596
QY 71 AGCCGGGAAAATCTCAGGAGTACAGTGCAGACCTTGAGGAGGCAACGCTCACTTGACCC 130
DB 597 AGCTGGAAAGCTCCGAGTTCACTGTGGCACCTGCAACCAAGCAACCTCACTTGAGCC 656
QY 131 AGGTCCATCTTCTGGAGTATGATGTTTAAATCCAAACCGGATGATGATGATGATGAT 190
DB 657 AGGTCCATCTTCTGGAGTATGATGTTTAAATCCAAACCGGATGATGATGATGATGAT 716
QY 191 CCCCACTGCTGGGAGTACAGTGCAGAAATTTTCTTAAATGAGACACCCCACT 250
DB 717 CTCGAGACTGCTGGGAGTACAGTGCAGAAATTTTCTTAAATGAGACACCCCACT 776
QY 251 CTGACAAAGAAACATCAGTACCTTTCAGCTGATGCAACAAATAGTGGAACTCACTT 310
DB 777 CAGACAAAGAACCTCGGTAGCTTGAACCTGATCACCAGCAAGGCGCAGCATCCCTT 836
QY 311 GCATTAGGTGACAGAGTGCAGAGCCCGTCTGTTTTCAGTGCAGACCTCCGCGACG 370
DB 837 GCATTAGGTGACAGAGTGCAGAGCCCGTCTGTTTTCAGTGCAGACCTCCGCGACG 896
QY 371 TGATTTGCTTACAGTGTTCCTTCACTTATGAGTGCAGAACTCAATGATGAGGAGTTTG 430
DB 897 TGATCTGTTTGAAGTGTTCCTTCACTTATGAGTGCAGAACTCAATGATGAGGAGTTTG 956
QY 431 TTCAGACCTCTCAACTTGGCTACTTCCCTGCTGTGT 467
DB 957 TTCAGATGCTCAACTTGGCTACTTCCCTGCTGTGT 993
RESULT 12
AAFS5259
ID AAF55259 standard; cDNA; 3255 BP.
XX AC AAF55259;
XX DT 29-MAY-2001 (first entry)
XX

DE Nucleotide sequence of a mutated murine parkin2 polypeptide.
 XX
 KM Parkinson's disease; 6q25.2-27; neurodegenerative disease;
 KM Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
 KM Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
 KM brain tumour; head trauma; stroke; vascular irregularity;
 KM metabolic irregularity; ss.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 129..1523
 FT /*tag= a
 FT /product= "truncated parkin2"
 XX
 XX EPI081225-A1.
 XX
 PD 07-MAR-2001.
 XX
 PF 30-AUG-1999; 99EP-00116766.
 XX
 PR 30-AUG-1999; 99EP-00116766.
 XX
 PA (BIOF-) BIOFRONTIERA PHARM GMBH.
 XX
 PI Luebbert H;
 XX
 XX WPI; 2001-212797/22.
 DR P-PSDB; AAB67532.
 XX
 XX New polynucleotides encoding mouse parkin2 protein, useful for producing
 PT a transgenic non-human animal as an animal model for neurodegenerative
 PT diseases.
 PT
 XX
 XX Claim 3; Page 36-38; 62pp; English.
 XX
 XX The present sequence encodes a murine parkin2 polypeptide. The sequence
 CC contains the mutation Thr145asn. Mutations or deletions in the parkin2
 CC gene cause Parkinson's disease in humans. The human parkin2 gene is
 CC located in gene region 6q25.2-27. parkin2 polypeptides and
 CC polynucleotides are useful for analysing neurodegenerative diseases. They
 CC are also useful for testing the efficacy of the treatment of a
 CC neurodegenerative disease such as Parkinson's disease, Alzheimer's
 CC disease, Huntington's disease, amyotrophic lateral sclerosis, Multi-
 CC system atrophy, Wilson's disease, Pick's disease, Prion disease, and
 CC secondary causes inducing Parkinson's syndromes like toxins, drugs, brain
 CC tumours, head trauma, stroke, vascular irregularities or metabolic
 CC irregularities, associated with a less active or non-active parkin
 CC protein.
 CC
 XX
 XX Sequence 3255 BP; 853 A; 817 C; 809 G; 776 T; 0 U; 0 Other;
 SQ
 Query Match 72.2%; Score 340.2; DB 4; Length 3255;
 Best Local Similarity 84.0%; Pred. No. 7.2e-107;
 Matches 384; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
 Oy 11 CAGTAGTCAATCATACAAAGCTTTTATGTGTATGCAAAAGCCCTGTAAAGAGTGC 70
 Db 538 CAGTTAAACCACTACAAAGCTTTTATCATCTAGTCAAAAGGCCCTGCCACAAGAGTCC 597
 Oy 71 AGCGGGGAAACTAGAGGTACAGTGCAGACCTGAGGAGCAACGCTCACTTGAACC 130
 Db 598 AGCGTGAAGACCTCGAGTTCAGTGTGACCTGCAAAAGCAACCTTCACTTGGGCC 657
 Oy 131 AGGGTCCATCTTGTGGGATGATGTTTAAATTCAAACCGGATGAGTGAATGCCAAT 190
 Db 658 AGGGCCCATCTTGTGGGAGCATGCTTAATTCAAACCGGATGAGTGAATGCCAAT 717
 Oy 191 CCCCACTGCTCCCTGGGAGTGTGCAAAATTTTCTTAAATGTGAGACAAACCCCACT 250
 Db 718 CTCGAGCTGCCCTGGAAACCAAGAGCTGAATTTTCTTAAATGTGAGACAAACCACT 777
 Oy 251 CTGACAGGAAGAAACATCAGTACTTGGACCTGATGCAAAAGTAATGTGAAACATCACTT 310

Db 778 CAGACAGGACACGTCGAGTCTTGAACCGATCACCAGCAAGGCGAGCATCCCTT 837
 Oy 311 GCATTATGTCACAGAGTACGAGGCCCGCCGTTTCCATGCAATCCGCCACG 370
 Db 838 GCATAGGCTGACAGAGTACGAGGCCCGCCGTTTCCATGATGTAACCAACCGTACG 897
 Oy 371 TGATTGCTTGAAGTCTTTCACCTTATATGATGCAAGACTCAATGATCGGAGTTTG 430
 Db 898 TGATCTGTTGAGACTGTTTCCACTTGTATGTGTGACAAAGCTCAACGATGCGAGTTTG 957
 Oy 431 TTCAGACCTCAACTTGGCTACTCCCTGCTGTGT 467
 Db 958 TCACAGTGTCAACTGCTACTCCCTGCGGTGTGT 994
 RESULT 13
 AAF55244
 ID AAF55244 standard; cDNA; 3255 BP.
 XX
 AC AAF55244;
 XX
 DT 29-MAY-2001 (first entry)
 XX
 XX Nucleotide sequence of a murine parkin2 polypeptide.
 XX
 KM Parkinson's disease; 6q25.2-27; neurodegenerative disease;
 KM Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
 KM Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease;
 KM brain tumour; head trauma; stroke; vascular irregularity;
 KM metabolic irregularity; ss.
 XX
 XX Mus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 129..1523
 FT /*tag= a
 FT /product= "parkin2"
 XX
 XX EPI081225-A1.
 XX
 PD 07-MAR-2001.
 XX
 PF 30-AUG-1999; 99EP-00116766.
 XX
 PR 30-AUG-1999; 99EP-00116766.
 XX
 PA (BIOF-) BIOFRONTIERA PHARM GMBH.
 XX
 PI Luebbert H;
 XX
 XX WPI; 2001-212797/22.
 DR P-PSDB; AAB67517.
 XX
 XX New polynucleotides encoding mouse parkin2 protein, useful for producing
 PT a transgenic non-human animal as an animal model for neurodegenerative
 PT diseases.
 PT
 XX
 XX Disclosure; Page 15-16; 62pp; English.
 XX
 XX The present sequence encodes a murine parkin2 polypeptide. Mutations or
 CC deletions in the parkin2 gene cause Parkinson's disease in humans. The
 CC human parkin2 gene is located in gene region 6q25.2-27. parkin2
 CC polypeptides and polynucleotides are useful for analysing
 CC neurodegenerative diseases. They are also useful for testing the efficacy
 CC of the treatment of a neurodegenerative disease such as Parkinson's
 CC disease, Alzheimer's disease, Huntington's disease, amyotrophic lateral
 CC sclerosis, Multi-system atrophy, Wilson's disease, Pick's disease, Prion
 CC disease, and secondary causes inducing Parkinson's syndromes like toxins,
 CC drugs, brain tumours, head trauma, stroke, vascular irregularities or
 CC metabolic irregularities, associated with a less active or non-active
 CC parkin protein

| Sequence | 32355 BP; | 852 A; | 818 C; | 809 G; | 776 T; | 0 U; | 0 Other; |
|-----------------------|---|---|--------|--------------|--------|--------|----------|
| Query Match | 72.2%; | Score 340.2; | DB 4; | Length 3255; | | | |
| Best Local Similarity | 84.0%; | Pred. No. 7.2e-107; | | | | | |
| Matches | 384; | Conservative | 0; | Mismatches | 73; | Indels | 0; |
| Db | | | | | | | |
| QY | 11 | CAGTAGATCATCTCAACACAGCTTTATGTGATTTGCAAGAGCCCTGTCAAAAGATGC | 70 | | | | |
| Db | 538 | CAGTTAAACCCACCTCAACACAGCTTTTATCTACTGCAAAAGCCCTGCAAGATCC | 597 | | | | |
| QY | 71 | AGCCGGAAAACTCAGGGTACAGTGCAGACCTGCAGGCGAAAGCTCACCTTGACCC | 130 | | | | |
| Db | 598 | AGCCTGGAAAGCTCCGAGTTCAGTGTGGCACCCTGCAAAACAGCAACCTCACCTTGACCC | 657 | | | | |
| QY | 131 | AGGGTCACCTCTGCTGGGATGTAGTTTAAATTCGAAACCGGATGAGTGTGAATGCCAAT | 190 | | | | |
| Db | 658 | AGGGCCCATCTGTGCTGGGAGCATGTCTTAATTCGAAACCGGATGAGTGTGATGCCAAT | 717 | | | | |
| QY | 191 | CCCCACACTGCCCTGGGACTAGTGCAGAAATTTTCTTTAAATGTGAGACACCCCACT | 250 | | | | |
| Db | 718 | CTCCAGACTGCCCTGGGAAACAGAGCTGAATTTTCTTTAAATGTGAGACACCCCACT | 777 | | | | |
| QY | 251 | CTGACAAAGAAACATCAGTAGCTTTGACACTGATGCCAACAATGTGGAAATCACTT | 310 | | | | |
| Db | 778 | CAGACAAAGAACGCTGGTAGCTTTGAACCTGATATCACAGCAACAGGGCGACATCCCTT | 837 | | | | |
| QY | 311 | GCATTACGTCACACAGCTCAGAGAGCCCGCTGCTGATTTTCCAGTGAATCCCGCACG | 370 | | | | |
| Db | 838 | GCATAGCGTCACACAGATGTCAGAGAGCCCTGCTGCTGCTCTTCAGATGTAACACCGTACG | 897 | | | | |
| QY | 371 | TGATTTGCTTAGACTGTGTTTCCACTTAACTGTGTGACAAAGACTCAATGATGGCAGTTTG | 430 | | | | |
| Db | 898 | TGATCTGTTTGGACGTTTCCACTGTGATTTGTGTGACAAAGACTCAACGATGGCAGTTTG | 957 | | | | |
| QY | 431 | TTACAGACCCCTCAACTTGGCTACTGCCGCTTGCT | 467 | | | | |
| Db | 958 | TCACGATGCTCACTTGGCTACTGCCGCTTGCT | 994 | | | | |
| RESULT 14 | | | | | | | |
| ID | AAFS5260 | | | | | | |
| XX | AAFS5260 | standard; cDNA; 3255 BP. | | | | | |
| XX | AAFS5260; | | | | | | |
| XX | AC | | | | | | |
| DT | 29-MAY-2001 | (first entry) | | | | | |
| XX | XX | | | | | | |
| DE | Nucleotide sequence of a mutated murine parkin2 polypeptide. | | | | | | |
| XX | XX | | | | | | |
| KW | Parkin2; Parkinson's disease; 6q25.2-27; neurodegenerative disease; | | | | | | |
| KW | Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis; | | | | | | |
| KW | Multi-system atrophy; Wilson's disease; Pick's disease; Prion disease; | | | | | | |
| KW | brain tumour; head trauma; stroke; vascular irregularity; | | | | | | |
| KW | metabolic irregularity; ss. | | | | | | |
| XX | XX | | | | | | |
| OS | Mus sp. | | | | | | |
| XX | XX | | | | | | |
| FT | Key | Location/Qualifiers | | | | | |
| FT | CDS | 129..1484 | | | | | |
| FT | FT | /tag= a | | | | | |
| FT | FT | /product= "truncated parkin2" | | | | | |
| XX | XX | | | | | | |
| PN | EP108125-A1. | | | | | | |
| PD | 07-MAR-2001. | | | | | | |
| XX | XX | | | | | | |
| PF | 30-AUG-1999; | 99EP-00116766. | | | | | |
| XX | XX | | | | | | |
| PR | 30-AUG-1999; | 99EP-00116766. | | | | | |
| XX | XX | | | | | | |
| PA | (BIOF-) BIOFRONTIERA PHARM GMBH. | | | | | | |
| XX | XX | | | | | | |
| PI | Luebbert H; | | | | | | |

| | |
|---------------------------|---|
| XX | WPI, 2001-212797/22. |
| DR | P-PDB; AAB67533. |
| XX | |
| PT | New polynucleotides encoding mouse parkin2 protein, useful for producing |
| PR | a transgenic non-human animal as an animal model for neurodegenerative |
| PT | diseases. |
| PS | |
| XX | Claim 3; Page 38-39; 62pp; English. |
| XX | |
| CC | The present sequence encodes a murine parkin2 polypeptide. The sequence |
| CC | contains the mutation Trp453stop. Mutations or deletions in the parkin2 |
| CC | gene cause Parkinson's disease in humans. The human parkin2 gene is |
| CC | located in gene region 6q25.2-27. Parkin2 polypeptides and |
| CC | polynucleotides are useful for analysing neurodegenerative diseases. They |
| CC | are also useful for testing the efficacy of the treatment of a |
| CC | neurodegenerative disease such as Parkinson's disease, Alzheimer's |
| CC | disease, Huntington's disease, amyotrophic lateral sclerosis, Multi- |
| CC | system atrophy, Wilson's disease, Pick's disease, prion disease, and |
| CC | secondary causes inducing Parkinson's syndromes like toxins, drugs, brain |
| CC | tumours, head trauma, stroke, vascular irregularities or metabolic |
| CC | irregularities, associated with a less active or non-active parkin |
| CC | protein |
| XX | |
| SQ | Sequence 3255 BP; 853 A; 818 C; 808 G; 776 T; 0 U; 0 Other; |
| | |
| Query Match | 72.2%; Score 340.2; DB 4; Length 3255; |
| Best Local Similarity | 84.0%; Pred No.7,2e-107; |
| Matches 384; Conservative | 0; Mismatches 73; Indels 0; Gaps 0 |
| OY | 11 CAGGTAGTCAATCTACAAACAGCTTTATGTAATTGGCAAGGCCCTGTCAAAGTGTC 70 |
| Db | 538 CAGTTAAACCACCTACAAACAGCTTTTCATCTA CTGCAAAAGGCCCTGSCACAAGTCC 597 |
| OY | 71 AGCCGGGAAAACTCAGGGTACAGTGCACACCTGGAAGCGACGACACTTGACCC 130 |
| Db | 598 AACCTGGAAAAGCTCCGAGTTCA GTGTGCACCTGCAACCAACCAACCTCACCTTGCCC 657 |
| OY | 131 AAGGTCCACTCTTGCTCGGAGTAGTATGTTTTAAATTCCAAACCGATGAGTGAATGCCAAT 190 |
| Db | 658 AGGGCCCATCTTGCTGGGACGATGCTTAATTCCAAACCGATGAGTGAATGCCAAGT 717 |
| OY | 191 CCCCACTGCCCCCTGGAGCTAGTGCAGAATTTTCTTTAAATGTGGAGCACACCCCACTT 250 |
| Db | 718 CTCACGATGCTCCCTGGAAACGAAGCTGAATTTTCTTTAAATGTGGAGCACACCCCACTT 777 |
| OY | 251 CTGACAAAGAAACATCAGTAGCTTGACCTGATGGCAACAAATAGTGGAAATCATCACTT 310 |
| Db | 778 CAGACAAAGAACACGTCGTAAGCTTTGAACCTGATACCAACGACAGGCGCAGATTCCTT 837 |
| OY | 311 GCATTAGCTGACACAGACGTCAGAGAGCCCGCTCTGCTGTTTTCAGTGCAACTCCCGCAAG 370 |
| Db | 838 GCATAGCTGACACAGATGTACAGAGAGCCCTGTCCTGAGTCTTCACAGTATACACACGTCACG 897 |
| OY | 371 TGATTGCTGTAGACGCTTTCCACTTATACGTGGAGCAAGACTCAATGATGGGCACTTGT 430 |
| Db | 898 TGATCTGTTTGACAGTGTTCACCTTGATGTGTGTCACAAAGACTCAAGATCGGAGGATTTG 957 |
| OY | 431 TTCAAGACCTCAACTTGCTGCTACTCTCCCTGCTTGCT 467 |
| Db | 958 TCACAGATGCTCAACTTGCTGCTACTCTCCCTGCTTGCT 994 |
| RESULT 15 | |
| AAF55258 | |
| ID | AAF55258 standard; cDNA; 3255 BP. |
| XX | |
| AC | AAF55258; |
| XX | |
| DT | 29-MAY-2001 (first entry) |
| XX | |
| DE | Nucleotide sequence of a mutated murine parkin2 polypeptide. |
| XX | |

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2005, 17:58:29 ; Search time 81.1144 Seconds
(without alignments)
9501.226 Million cell updates/sec

Title: US-10-622-817-10

Perfect score: 471
Sequence: 1 ggaagaccagcagtagatc.....acccctccttggtgtag 471

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 81813859 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/lna/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/lna/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/lna/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/lna/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/lna/PTCTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/lna/backfilest.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|-------------------|
| 1 | 466.4 | 99.0 | 2960 | US-09-601-844B-1 | Sequence 1, Appli |
| 2 | 466.4 | 99.0 | 2960 | US-09-949-016-645 | Sequence 645, App |
| 3 | 288.4 | 61.2 | 2876 | US-09-601-844B-3 | Sequence 3, Appli |
| 4 | 288.4 | 61.2 | 2876 | US-09-949-016-982 | Sequence 982, App |
| 5 | 251.8 | 53.5 | 2513 | US-09-949-016-983 | Sequence 983, App |
| 6 | 37.8 | 8.0 | 399 | US-09-621-976-8976 | Sequence 8976, Ap |
| 7 | 34.8 | 7.4 | 40 | US-09-601-844B-65 | Sequence 65, Appl |
| 8 | 34.8 | 7.4 | 361 | US-09-643-597-293 | Sequence 293, App |
| 9 | 34.8 | 7.4 | 361 | US-09-480-884A-293 | Sequence 293, App |
| 10 | 34.8 | 7.4 | 361 | US-09-542-615A-293 | Sequence 293, App |
| 11 | 34.8 | 7.4 | 361 | US-09-606-421B-293 | Sequence 293, App |
| 12 | 34.8 | 7.4 | 361 | US-09-630-940B-293 | Sequence 293, App |
| 13 | 34.8 | 7.4 | 405 | US-09-513-999C-11324 | Sequence 11324, A |
| 14 | 34.8 | 7.2 | 40 | US-09-601-844B-64 | Sequence 64, Appl |
| 15 | 32.8 | 7.0 | 178883 | US-09-949-016-12733 | Sequence 12733, A |
| 16 | 32.8 | 7.0 | 178884 | US-09-949-016-13039 | Sequence 13039, A |
| 17 | 32.6 | 6.9 | 35262 | US-09-949-016-12797 | Sequence 12797, A |
| 18 | 32.6 | 6.9 | 35263 | US-09-949-016-16399 | Sequence 16399, A |
| 19 | 32.2 | 6.8 | 152132 | US-09-949-016-13845 | Sequence 13845, A |
| 20 | 32.2 | 6.8 | 152145 | US-09-949-016-12371 | Sequence 12371, A |
| 21 | 32.2 | 6.8 | 38 | US-09-601-844B-66 | Sequence 66, Appl |
| 22 | 32.2 | 6.8 | 347 | US-09-621-976-11665 | Sequence 11665, A |
| 23 | 32.2 | 6.8 | 819 | US-09-489-039A-3239 | Sequence 3239, Ap |
| 24 | 31.8 | 6.8 | 37 | US-09-601-844B-68 | Sequence 68, Appl |
| 25 | 31.8 | 6.8 | 342 | US-09-134-000C-2019 | Sequence 2019, Ap |
| 26 | 31.8 | 6.8 | 1160 | US-09-737-698B-12 | Sequence 12, Appl |
| 27 | 31.8 | 6.8 | 1160 | US-09-737-626A-12 | Sequence 12, Appl |

| | | | | | |
|----|------|-----|---------|----------------------|--------------------|
| 28 | 31.8 | 6.8 | 1695 | US-09-737-698B-28 | Sequence 28, Appl |
| 29 | 31.8 | 6.8 | 1695 | US-09-737-626A-28 | Sequence 28, Appl |
| 30 | 31.8 | 6.8 | 1830121 | US-09-557-884-1 | Sequence 1, Appl1 |
| 31 | 31.8 | 6.8 | 1830121 | US-09-643-990A-1 | Sequence 1, Appl1 |
| 32 | 31.6 | 6.7 | 601 | US-09-949-016-36816 | Sequence 36816, A |
| 33 | 31.6 | 6.7 | 601 | US-09-949-016-58953 | Sequence 58953, A |
| 34 | 31.6 | 6.7 | 601 | US-09-949-016-184559 | Sequence 184559, A |
| 35 | 31.6 | 6.7 | 119801 | US-09-949-016-13453 | Sequence 13453, A |
| 36 | 31.6 | 6.7 | 190078 | US-09-949-016-12507 | Sequence 12507, A |
| 37 | 31.6 | 6.7 | 190078 | US-09-949-016-12707 | Sequence 12707, A |
| 38 | 31.6 | 6.7 | 190078 | US-09-949-016-17026 | Sequence 17026, A |
| 39 | 31.6 | 6.7 | 336024 | US-09-949-016-14043 | Sequence 14043, A |
| 40 | 31.4 | 6.7 | 336024 | US-09-949-016-12373 | Sequence 12373, A |
| 41 | 31.2 | 6.6 | 832 | US-09-621-976-2813 | Sequence 2813, Ap |
| 42 | 31.2 | 6.6 | 27727 | US-09-949-016-15737 | Sequence 15737, A |
| 43 | 31.2 | 6.6 | 29067 | US-09-949-016-12535 | Sequence 12535, A |
| 44 | 31.2 | 6.6 | 29093 | US-09-949-016-17476 | Sequence 17476, A |
| 45 | 30.8 | 6.5 | 7194 | US-09-949-016-17512 | Sequence 17512, A |
| 45 | 30.8 | 6.5 | 7194 | US-09-601-326-76 | Sequence 76, Appl |

ALIGNMENTS

RESULT 1
US-09-601-844B-1
; Sequence 1, Application US/09601844B
; Patent No. 6716621
; GENERAL INFORMATION:
; APPLICANT: Shimizu, No. 6716621yoshi
; APPLICANT: Mizuno, Yoshikuni
; TITLE OF INVENTION: Isolated DNA or Gene Responsible for Parkinson's Disease
; FILE REFERENCE: 0652 211000
; CURRENT APPLICATION NUMBER: US/09/601, 844B
; CURRENT FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: PCT/JP99/00545
; PRIOR FILING DATE: 1999-02-09
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2960
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (102)..(1496)
; FEATURE:
; OTHER INFORMATION:
; NAME/KEY: misc feature
; LOCATION: (109)..(272)
; OTHER INFORMATION: Exon 1
; NAME/KEY: misc feature
; LOCATION: (109)..(272)
; OTHER INFORMATION: Exon 2
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (273)..(513)
; OTHER INFORMATION: Exon 3
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (514)..(635)
; OTHER INFORMATION: Exon 4
; NAME/KEY: misc feature
; LOCATION: (636)..(719)
; OTHER INFORMATION: Exon 5
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (720)..(835)
; OTHER INFORMATION: Exon 6
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (836)..(972)

```
/ OTHER INFORMATION: Exon 7
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (973)..(1034)
/ OTHER INFORMATION: Exon 8
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1035)..(1184)
/ OTHER INFORMATION: Exon 9
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1185)..(1268)
/ OTHER INFORMATION: Exon 10
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1269)..(1386)
/ OTHER INFORMATION: Exon 11
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1387)..(2960)
/ OTHER INFORMATION: Exon 12
US-09-601-844B-1
```

```
Query Match      99.0%; Score 466.4; DB 4; Length 2960;
Best Local Similarity 99.8%; Pred. No. 1,9e-154;
Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 GGAAGTCCAGAGGTAGATCAATCTACAAAGCTTTTATGTATTTGAAAGGCCCTGT 60
DB 504 GGAAGTCCAGAGGTAGATCAATCTACAAAGCTTTTATGTATTTGAAAGGCCCTGT 563
QY 61 CAAGAAGTGCAGCCGGGAAAACCTCAGGGTACAGTGCAGACCTGCAGGCAAGCGCTC 120
DB 564 CAAGAAGTGCAGCCGGGAAAACCTCAGGGTACAGTGCAGACCTGCAGGCAAGCGCTC 623
QY 121 ACCTTGACCCAGGGTCCATCTTGTGAGATGATGTTTAAATCCAAACCGATGAGTGT 180
DB 624 ACCTTGACCCAGGGTCCATCTTGTGAGATGATGTTTAAATCCAAACCGATGAGTGT 683
QY 181 GAATGCCAATCCCAACCTGCGCTGAGACTAGTGCAGAAATTTTCTTAAATGTGAGCA 240
DB 684 GAATGCCAATCCCAACCTGCGCTGAGACTAGTGCAGAAATTTTCTTAAATGTGAGCA 743
QY 241 CACCCCACTCTGACAGAAACATCAGTAGCTTTGACCTGATCGCAACAAATAGTCGG 300
DB 744 CACCCCACTCTGACAGAAACATCAGTAGCTTTGACCTGATCGCAACAAATAGTCGG 803
QY 301 AACATCACTTGATTAAGTGCAGAGCGTCAAGAGCCCGCTGCTGTTTCCAGTGCAC 360
DB 804 AACATCACTTGATTAAGTGCAGAGCGTCAAGAGCCCGCTGCTGTTTCCAGTGCAC 863
QY 361 TCCCGCAGCTGATTTGCTTACAGCTGTTCCACTTATCTGTGACAGACATCATGAT 420
DB 864 TCCCGCAGCTGATTTGCTTACAGCTGTTCCACTTATCTGTGACAGACATCATGAT 923
QY 421 CGGAGATTGTTTCAGACACCTCAACTGGCTACTCCCTGCTGTGTG 468
DB 924 CGGAGATTGTTTCAGACACCTCAACTGGCTACTCCCTGCTGTGTG 971
```

```
RESULT 2
US-09-949-016-645
Sequence 645, Application US/09949016
Patent No. 681239
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
```

```
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 645
/ LENGTH: 2960
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-645
```

```
Query Match      99.0%; Score 466.4; DB 4; Length 2960;
Best Local Similarity 99.8%; Pred. No. 1,9e-154;
Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 GGAAGTCCAGAGGTAGATCAATCTACAAAGCTTTTATGTATTTGAAAGGCCCTGT 60
DB 504 GGAAGTCCAGAGGTAGATCAATCTACAAAGCTTTTATGTATTTGAAAGGCCCTGT 563
QY 61 CAAGAAGTGCAGCCGGGAAAACCTCAGGGTACAGTGCAGACCTGCAGGCAAGCGCTC 120
DB 564 CAAGAAGTGCAGCCGGGAAAACCTCAGGGTACAGTGCAGACCTGCAGGCAAGCGCTC 623
QY 121 ACCTTGACCCAGGGTCCATCTTGTGAGATGATGTTTAAATCCAAACCGATGAGTGT 180
DB 624 ACCTTGACCCAGGGTCCATCTTGTGAGATGATGTTTAAATCCAAACCGATGAGTGT 683
QY 181 GAATGCCAATCCCAACCTGCGCTGAGACTAGTGCAGAAATTTTCTTAAATGTGAGCA 240
DB 684 GAATGCCAATCCCAACCTGCGCTGAGACTAGTGCAGAAATTTTCTTAAATGTGAGCA 743
QY 241 CACCCCACTCTGACAGAAACATCAGTAGCTTTGACCTGATCGCAACAAATAGTCGG 300
DB 744 CACCCCACTCTGACAGAAACATCAGTAGCTTTGACCTGATCGCAACAAATAGTCGG 803
QY 301 AACATCACTTGATTAAGTGCAGAGCGTCAAGAGCCCGCTGCTGTTTCCAGTGCAC 360
DB 804 AACATCACTTGATTAAGTGCAGAGCGTCAAGAGCCCGCTGCTGTTTCCAGTGCAC 863
QY 361 TCCCGCAGCTGATTTGCTTACAGCTGTTCCACTTATCTGTGACAGACATCATGAT 420
DB 864 TCCCGCAGCTGATTTGCTTACAGCTGTTCCACTTATCTGTGACAGACATCATGAT 923
QY 421 CGGAGATTGTTTCAGACACCTCAACTGGCTACTCCCTGCTGTGTG 468
DB 924 CGGAGATTGTTTCAGACACCTCAACTGGCTACTCCCTGCTGTGTG 971
```

```
RESULT 3
US-09-601-844B-3
Sequence 3, Application US/09601844B
Patent No. 6716621
GENERAL INFORMATION:
APPLICANT: Shimizu, No. 6716621uyoshi
TITLE OF INVENTION: Isolated DNA or Gene Responsible for Parkinson's Disease
FILE REFERENCE: 0652,211000
CURRENT APPLICATION NUMBER: US/09/601,844B
PRIOR FILING DATE: 2000-08-09
PRIOR APPLICATION NUMBER: PCT/jp99/00545
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn version 3.1.1
SEQ ID NO 3
LENGTH: 2876
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (102)..(1412)
OTHER INFORMATION:
FEATURE:
```

NAME/KEY: misc feature
LOCATION: (102)..(108)
OTHER INFORMATION: Exon 1
FEATURE:
NAME/KEY: misc feature
LOCATION: (109)..(272)
OTHER INFORMATION: Exon 2
FEATURE:
NAME/KEY: misc feature
LOCATION: (273)..(513)
OTHER INFORMATION: Exon 3
FEATURE:
NAME/KEY: misc feature
LOCATION: (514)..(635)
OTHER INFORMATION: Exon 4
FEATURE:
NAME/KEY: misc feature
LOCATION: (636)..(751)
OTHER INFORMATION: Exon 6
FEATURE:
NAME/KEY: misc feature
LOCATION: (752)..(888)
OTHER INFORMATION: Exon 7
FEATURE:
NAME/KEY: misc feature
LOCATION: (889)..(950)
OTHER INFORMATION: Exon 8
FEATURE:
NAME/KEY: misc feature
LOCATION: (951)..(1100)
OTHER INFORMATION: Exon 9
FEATURE:
NAME/KEY: misc feature
LOCATION: (1101)..(1184)
OTHER INFORMATION: Exon 10
FEATURE:
NAME/KEY: misc feature
LOCATION: (1185)..(1302)
OTHER INFORMATION: Exon 11
FEATURE:
NAME/KEY: misc feature
LOCATION: (1303)..(2876)
OTHER INFORMATION: Exon 11
US-09-601-844B-3

Query Match 61.2%; Score 288.4; DB 4; Length 2876;
Best local Similarity 81.8%; Pred. No. 2.7e-91;
Matches 383; Conservative 0; Mismatches 1; Indels 84; Gaps 1;

1 GGAAGTCCAGAGGTAAGTCAATCTACAACAGCTTTTATGTGATTTGCAAAAGGCCCTGT 60
504 GGAAGTCCAGAGGTAAGTCAATCTACAACAGCTTTTATGTGATTTGCAAAAGGCCCTGT 563
61 CAAGAGTGCAGCCGGGAAAACCTCAGGGTACAGTGCAGCACTGCAGGCAAGCGCTC 120
564 CAAGAGTGCAGCCGGGAAAACCTCAGGGTACAGTGCAGCACTGCAGGCAAGCGCTC 623
121 ACCTTGACCCAGGGTCCATCTTGCTGGAGTATGTTTAATTCAAACCGGATGATGCT 180
624 ACCTTGACCCAG-----GAAATTTCTTTAAATGTGAGCA 635
181 GAATGCCAATCCCAACACTGCGCTGGACTAGTGAGAAATTTTCTTTAAATGTGAGCA 240
636 -----GAAATTTCTTTAAATGTGAGCA 659
241 CACCCCACTCTGCACAAGAAACATCAGTAGCTTTGCACTGATCGCAACAAATAGTCGG 300
660 CACCCCACTCTGCACAAGAAACATCAGTAGCTTTGCACTGATCGCAACAAATAGTCGG 719
301 AACATCACTTGCAATTAAGTGCAGACAGCTCAGAGAGCCCGTCTGCTTTTCCAGTGCAAC 360
720 AACATCACTTGCAATTAAGTGCAGACAGCTCAGAGAGCCCGTCTGCTTTTCCAGTGCAAC 779

361 TCCCGCAGCGATTTGCTTAGACTGTTTCCATTATAGTGTGACAAAGACTCAATGAT 420
780 TCCCGCAGCGATTTGCTTAGACTGTTTCCATTATAGTGTGACAAAGACTCAATGAT 839
421 CGGCAATTTGTTACGACACCTCAACTGGCTACTCCCTGCTTGTG 468
840 CGGCAATTTGTTACGACACCTCAACTGGCTACTCCCTGCTTGTG 887

RESULT 4
US-09-949-016-982
Sequence 982, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT FILING DATE: 2000-04-14
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 982
LENGTH: 2876
TYPE: DNA
ORGANISM: Human
US-09-949-016-982

Query Match 61.2%; Score 288.4; DB 4; Length 2876;
Best local Similarity 81.8%; Pred. No. 2.7e-91;
Matches 383; Conservative 0; Mismatches 1; Indels 84; Gaps 1;

1 GGAAGTCCAGAGGTAAGTCAATCTACAACAGCTTTTATGTGATTTGCAAAAGGCCCTGT 60
504 GGAAGTCCAGAGGTAAGTCAATCTACAACAGCTTTTATGTGATTTGCAAAAGGCCCTGT 563
61 CAAGAGTGCAGCCGGGAAAACCTCAGGGTACAGTGCAGCACTGCAGGCAAGCGCTC 120
564 CAAGAGTGCAGCCGGGAAAACCTCAGGGTACAGTGCAGCACTGCAGGCAAGCGCTC 623
121 ACCTTGACCCAGGGTCCATCTTGCTGGAGTATGTTTAATTCAAACCGGATGATGCT 180
624 ACCTTGACCCAG-----GAAATTTCTTTAAATGTGAGCA 635
181 GAATGCCAATCCCAACACTGCGCTGGACTAGTGAGAAATTTTCTTTAAATGTGAGCA 240
636 -----GAAATTTCTTTAAATGTGAGCA 659
241 CACCCCACTCTGCACAAGAAACATCAGTAGCTTTGCACTGATCGCAACAAATAGTCGG 300
660 CACCCCACTCTGCACAAGAAACATCAGTAGCTTTGCACTGATCGCAACAAATAGTCGG 719
301 AACATCACTTGCAATTAAGTGCAGACAGCTCAGAGAGCCCGTCTGCTTTTCCAGTGCAAC 360
720 AACATCACTTGCAATTAAGTGCAGACAGCTCAGAGAGCCCGTCTGCTTTTCCAGTGCAAC 779
361 TCCCGCAGCGATTTGCTTAGACTGTTTCCATTATAGTGTGACAAAGACTCAATGAT 420
780 TCCCGCAGCGATTTGCTTAGACTGTTTCCATTATAGTGTGACAAAGACTCAATGAT 839
421 CGGCAATTTGTTACGACACCTCAACTGGCTACTCCCTGCTTGTG 468
840 CGGCAATTTGTTACGACACCTCAACTGGCTACTCCCTGCTTGTG 887

RESULT 5
US-09-949-016-983

```
/ Sequence 983, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 983
/ LENGTH: 2513
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-983
```

```
Query Match      53.5%; Score 251.8; DB 4; Length 2513;
Best Local Similarity 97.3%; Pred. No. 2.4e-78;
Matches 256; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

```
QY 206 GCACTAGTGCAGAAATTTTCTTAATGTGAGACACCCCACTTCGACGAAGAAT 265
DB 262 GCAGCTGCAGAAATTTTCTTAATGTGAGACACCCCACTTCGACGAAGAAT 321
QY 266 CAGTACCTTTCAGCTGATCGCAACAAATAGTCGGAACATCATCTGATTAAGTCA 325
DB 322 CAGTACCTTTCAGCTGATCGCAACAAATAGTCGGAACATCATCTGATTAAGTCA 381
QY 326 ACGTACGAGCCCGCTCTGTTTCCAGTGCACATCTCCCGCAGTATTTGCTTAAGT 385
DB 382 ACGTACGAGCCCGCTCTGTTTCCAGTGCACATCTCCCGCAGTATTTGCTTAAGT 441
QY 386 GTTTCACCTTATCTGTGTGACAGACTCATGTGAGCTGTTGTTCACGACCTCAAC 445
DB 442 GTTTCACCTTATCTGTGTGACAGACTCATGTGAGCTGTTGTTCACGACCTCAAC 501
QY 446 TTGGCTACTCCCTGCTTGTGTG 468
DB 502 TTGGCTACTCCCTGCTTGTGTG 524
```

```
RESULT 6
US-09-621-976-8976
/ Sequence 8976, Application US/09621976
/ Patent No. 6639063
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Dumas Milne Edwards, J.B.
/ APPLICANT: Jobert, S.
/ APPLICANT: Giordano, J.Y.
/ TITLE OF INVENTION: ESTs and Encoded Human Proteins.
/ FILE REFERENCE: GENSER.054PR2
/ CURRENT APPLICATION NUMBER: US/09/621,976
/ CURRENT FILING DATE: 2000-07-21
/ NUMBER OF SEQ ID NOS: 19335
/ SOFTWARE: Patent .pm
/ SEQ ID NO: 8976
/ LENGTH: 399
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-621-976-8976
```

```
Query Match      8.0%; Score 37.8; DB 4; Length 399;
Best Local Similarity 12.0%; Pred. No. 0.0069;
Matches 36; Conservative 13; Mismatches 11; Indels 0; Gaps 0;
```

```
QY 166 AACCGATGATGTGTGAATGCCAATCCCACTGCTCCCTGGACTGTGCAGAATTTTC 225
```

```
DB 20 RAYGMYTTRGSKSRKRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 79
QY 226 TTTAAATGTGAGACACCCCACTTCGACGAAGAATCATGATGCTTTCACCTGATC 285
DB 80 RGAASWAGYMSWMTYRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 139
QY 286 GCAACAATATGTGGACATCATCTGATTAAGTGCACAGACGTGACGACCCCGCTCTG 345
DB 140 AMKTMAMKSKSRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 199
QY 346 GTTTCAGTGCACATCCCGCAGTATTTGCTTACATGTTTCCACTTATCTGTGTG 405
DB 200 RKKMYGASGMSWRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 259
QY 406 ACAAGATCATGATGCGGAGTTTGTTCACGACCCCTCAATGCTGCTACTCCCTGCT 465
DB 260 CCCMSGSCWTKTYRRSWYCASCYYSYKTRASCMMCCMKRRMMAMMYMMYTCCKT 319
QY 466 G 466
DB 320 S 320
```

```
RESULT 7
US-09-601-844B-65
/ Sequence 65, Application US/09601844B
/ Patent No. 6716621
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Shimizu, No. 6716621uyoshi
/ APPLICANT: Mizuno, Yoshihumi
/ TITLE OF INVENTION: Isolated DNA or Gene Responsible for Parkinson's Disease
/ FILE REFERENCE: 0652-2110000
/ CURRENT APPLICATION NUMBER: US/09/601,844B
/ PRIOR FILING DATE: 2000-08-09
/ PRIOR APPLICATION NUMBER: PCT/JP99/00545
/ PRIOR FILING DATE: 1999-02-09
/ NUMBER OF SEQ ID NOS: 70
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO: 65
/ LENGTH: 40
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-601-844B-65
```

```
Query Match      7.4%; Score 34.8; DB 4; Length 40;
Best Local Similarity 94.7%; Pred. No. 0.02;
Matches 36; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 128 CCCAGGTCATCTGCTGGATGATGTTTAAATTC 165
DB 3 CCAAGTGCATCTGCTGGATGATGTTTAAATTC 40
```

```
RESULT 8
US-09-643-597-293
/ Sequence 293, Application US/09643597
/ Patent No. 6426072
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Wang, Tongrong
/ APPLICANT: Fan, Liqun
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Hosken, Nancy
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Li, Samuel X.
/ APPLICANT: Wang, Aijun
/ APPLICANT: Skeiky, Yaelir A.W.
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: McNeill, Patricia D.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.455C11
```


;; CURRENT APPLICATION NUMBER: US/09/643,597
;; CURRENT FILING DATE: 2000-08-21
;; NUMBER OF SEQ ID NOS: 369
;; SOFTWARE: FASTSEQ for Windows Version 3.0
;; SEQ ID NO 293
;; LENGTH: 361
;; TYPE: DNA
;; ORGANISM: Homo sapien
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(361)
;; OTHER INFORMATION: n = A,T,C or G
US-09-643-597-293

Query Match 7.4%; Score 34.8; DB 3; Length 361;
Best Local Similarity 59.4%; Pred.No.0.076;
Matches 57; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 203 CTGGAGCTAGTGCAGAAATTTTCTTTAAATGTGAGACACCCCACTCTGACAAGAA 262
DB 134 CTTGAACACGCTTCAATTTTCTTCCAAATGCTGACATGCCACCTTGAGTAAGAA 193
QY 263 CATCAGTAGCTTTGACCTGATGCGACAAATATGTC 298
DB 194 CANAAGTATTTTAAACATGACAGCTAANNAACATTC 229

RESULT 9
US-09-480-884A-293
; Sequence 293, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 293
; LENGTH: 361
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(361)
; OTHER INFORMATION: n = A,T,C or G
US-09-480-884A-293

Query Match 7.4%; Score 34.8; DB 4; Length 361;
Best Local Similarity 59.4%; Pred.No.0.076;
Matches 57; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 203 CTGGAGCTAGTGCAGAAATTTTCTTTAAATGTGAGACACCCCACTCTGACAAGAA 262
DB 134 CTTGAACACGCTTCAATTTTCTTCCAAATGCTGACATGCCACCTTGAGTAAGAA 193
QY 263 CATCAGTAGCTTTGACCTGATGCGACAAATATGTC 298
DB 194 CANAAGTATTTTAAACATGACAGCTAANNAACATTC 229

RESULT 10
US-09-542-615A-293
; Sequence 293, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong

;; APPLICANT: Fan, Liqun
;; APPLICANT: Kalos, Michael D.
;; APPLICANT: Bangur, Chaitanya S.
;; APPLICANT: Hosken, Nancy A.
;; APPLICANT: Fanger, Gary R.
;; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
;; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;; FILE REFERENCE: 210121.455C8
;; CURRENT APPLICATION NUMBER: US/09/542,615A
;; CURRENT FILING DATE: 2000-04-14
;; NUMBER OF SEQ ID NOS: 350
;; SOFTWARE: FASTSEQ for Windows Version 3.0
;; SEQ ID NO 293
;; LENGTH: 361
;; TYPE: DNA
;; ORGANISM: Homo sapien
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(361)
;; OTHER INFORMATION: n = A,T,C or G
US-09-542-615A-293

Query Match 7.4%; Score 34.8; DB 4; Length 361;
Best Local Similarity 59.4%; Pred.No.0.076;
Matches 57; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 203 CTGGAGCTAGTGCAGAAATTTTCTTTAAATGTGAGACACCCCACTCTGACAAGAA 262
DB 134 CTTGAACACGCTTCAATTTTCTTCCAAATGCTGACATGCCACCTTGAGTAAGAA 193
QY 263 CATCAGTAGCTTTGACCTGATGCGACAAATATGTC 298
DB 194 CANAAGTATTTTAAACATGACAGCTAANNAACATTC 229

RESULT 11
US-09-606-421B-293
; Sequence 293, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skelky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 293
; LENGTH: 361
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(361)
; OTHER INFORMATION: n = A,T,C or G
US-09-606-421B-293

Query Match 7.4%; Score 34.8; DB 4; Length 361;
Best Local Similarity 59.4%; Pred.No.0.076;
Matches 57; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 203 CTGGAGCTAGTGCAGAAATTTTCTTTAAATGTGAGACACCCCACTCTGACAAGAA 262
DB 134 CTTGAACACGCTTCAATTTTCTTCCAAATGCTGACATGCCACCTTGAGTAAGAA 193

Qy 263 CATCAGTAGCTTTGCACCTGATCGCAACAATAAGTC 298
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 194 CANAAGTATTTTAACATGACACAGCTAANAAACATTC 229

```

US-09-630-940B-293
RESULT 12
US-09-630-940B-293
Sequence 293 Application US/09630940B
Patent No. 6737514
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Fan, Liqun
APPLICANT: Kalos, Michael D.
APPLICANT: Bangur, Chaltanya S.
APPLICANT: Hosken, Nancy
APPLICANT: Pangser, Gary R.
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skelky, Yasir A.W.
APPLICANT: Henderson, Robert A.
APPLICANT: McNeill, Patricia D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPEUTIC TREATMENT OF LUNG CANCER
FILE REFERENCE: 210121.455C10
CURRENT APPLICATION NUMBER: US/09/630,940B
CURRENT FILING DATE: 2000-08-02
NUMBER OF SEQ ID NOS: 367
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 293
LENGTH: 361
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(361)
OTHER INFORMATION: n = A,T,C or G
US-09-630-940B-293

```

| | Query Match | Similarity | Score | DB | Length |
|------------|-------------|--------------|-------|------------|--------|
| Best Local | 59.4% | | 34.8 | DB 4 | 361 |
| Matches | 57 | Conservative | 0 | Mismatches | 39 |
| | | | | Indels | 0 |
| | | | | Gaps | 0 |

```

RESULT 13
US-09-513-999C-11324/C
; Sequence 11324, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J. B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J. Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 11324
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Homo sapiens

```

```

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 326
; OTHER INFORMATION: w=a or t
US-09-513-999C-11324

```

| | | | | |
|-----------------------|-----------------|------------------|-----------|-------------|
| Query Match | 7.4% | Score 34.8; | DB 4; | Length 405; |
| Best Local Similarity | 59.4%; | Pred. No. 0.081; | | |
| Matches 57; | Conservative 1; | Mismatches 38; | Indels 0; | Gaps 0; |

QY 203 CTGGAGCTAATGAGAAATTTTCTTTAAATGAGAGACACCCACCTGTACAAAGAAA 262
 Db 333 CTTGAACWCTCTTCAATTTCTTCCCAAANCGTGCATGCACTTGAAGTAACGAAG 274
 QY 263 CATCAGTAGCTTTGCACCTGATCGCAAAATATGTC 298
 Db 273 CAGAGATATTTTAAACATGACACGCTAAGAACATTTC 238

```

RESULT 14
US-09-601-844B-64
; Sequence 64, Application US/09601844B
; Patent No. 6716621
; GENERAL INFORMATION:
; APPLICANT: Shimizu, Yoshiaki
; TITLE OF INVENTION: Isolated DNA or Gene Responsible for Parkinson's Disease
; FILE REFERENCE: 0652.211000
; CURRENT APPLICATION NUMBER: US/09/601.844B
; CURRENT FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: PCT/JP99/00545
; PRIOR FILING DATE: 1999-02-09
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-601-844B-64

```

| | | | | |
|-------------|--------|-----------------------------------|--------------|-----------|
| Query Match | 7.2% | Score 34 | DB 4 | Length 40 |
| Similarity | 100.0% | Pred. No. 0.038 | | |
| Local | | | | |
| Matches | 34 | Conservative 0 | Mismatches 0 | Indels 0 |
| | | | | |
| QY | 131 | AGGGTCATCTTGGCTGGGATGATGTTTAAATCC | 164 | |
| | | | | |
| Db | 7 | AGGGTCATCTTGGCTGGGATGATGTTTAAATCC | 40 | |

```

US-09-949-016-12733/C
Sequence 12733, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12733
LENGTH: 178883
TYPE: DNA
ORGANISM: Human
US-09-949-016-12733

```

Query Match 7.0%; Score 32.8; DB 4; Length 178883;
Best Local Similarity 54.0%; Pred. No. 17;
Matches 67; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

| | | | |
|----|--------|--|--------|
| Qy | 33 | CTTTATGTTGTAATGCAAGGCCCTGTCAAGAGTGCAGCCGGGAAACTCAGGGTACA | 92 |
| Db | 109301 | CTTGTCGGTGTGATGCAATGTGCGGTTTGCAAAATGTGACTCACAGCCAGCGGCTTACA | 109242 |
| Qy | 93 | GTGCAGCACTGCAGGCGGCAAGCTCACCCTTGACCCAGGGTCCATCTTGCTGGGATGA | 152 |
| Db | 109241 | AACCTAGTGTCTGCAGGAGGAACAGTCATCCAGAAATCAGGGCACAGTTGCCGTTGGTGC | 109182 |
| Qy | 153 | TGTT | 156 |
| Db | 109181 | TGCT | 109178 |

Search completed: February 26, 2005, 05:20:30
Job time : 88.114 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using BW model

Run on: February 26, 2005, 02:13:05 ; Search time 283.738 Seconds
(without alignments)
9836.161 Million cell updates/sec

Title: US-10-622-817-10

Perfect score: 471

Sequence: 1 ggaagtcacagcagctagatc.....actccctgctgtgtgttag 471

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5394803 seqs, 2962729879 residues

Total number of hits satisfying chosen parameters: 10789606

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*

20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*

21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|------------------|
| 1 | 471 | 100.0 | 471 | 9 | US-09-785-548-3 |
| 2 | 471 | 100.0 | 471 | 18 | US-10-622-817-10 |
| 3 | 466.4 | 99.0 | 2955 | 19 | US-10-839-688-12 |
| 4 | 466.4 | 99.0 | 2960 | 18 | US-10-473-226-1 |
| 5 | 466.4 | 99.0 | 2960 | 18 | US-10-473-226-3 |
| 6 | 466.4 | 99.0 | 2960 | 18 | US-10-622-817-7 |
| 7 | 466.4 | 99.0 | 2960 | 18 | US-10-776-604-1 |
| 8 | 466.4 | 99.0 | 2960 | 18 | US-10-839-688-11 |
| 9 | 347 | 73.7 | 1470 | 19 | US-10-622-817-13 |
| 10 | 288.4 | 61.2 | 2876 | 18 | US-10-776-604-3 |
| 11 | 138.6 | 29.4 | 650 | 19 | US-10-839-688-6 |

| | 12 | 89.8 | 19.1 | 660 | 19 | US-10-839-688-4 | Sequence 4, Appli |
|--|----|------|------|---------|----|----------------------|--------------------|
| | 13 | 50.2 | 10.7 | 45 | 10 | US-09-908-975-26043 | Sequence 26043, A |
| | 14 | 34.8 | 7.4 | 60 | 18 | US-10-776-604-65 | Sequence 65, Appl |
| | 15 | 34.8 | 7.4 | 361 | 9 | US-09-735-705-293 | Sequence 293, App |
| | 16 | 34.8 | 7.4 | 361 | 9 | US-09-830-716A-293 | Sequence 293, App |
| | 17 | 34.8 | 7.4 | 361 | 9 | US-09-897-778-293 | Sequence 293, App |
| | 18 | 34.8 | 7.4 | 361 | 14 | US-10-007-700-293 | Sequence 293, App |
| | 19 | 34.8 | 7.4 | 361 | 15 | US-10-117-982-293 | Sequence 293, App |
| | 20 | 34.8 | 7.4 | 361 | 17 | US-10-313-986-293 | Sequence 293, App |
| | 21 | 34.8 | 7.4 | 361 | 18 | US-10-775-972-293 | Sequence 293, App |
| | 22 | 34.6 | 7.3 | 339 | 19 | US-10-696-639-819 | Sequence 819, App |
| | 23 | 34 | 7.2 | 40 | 18 | US-10-776-604-64 | Sequence 64, Appl |
| | 24 | 33.6 | 7.1 | 203 | 17 | US-10-242-535A-10255 | Sequence 10255, A |
| | 25 | 33.6 | 7.1 | 203 | 17 | US-10-085-783A-10255 | Sequence 10255, A |
| | 26 | 33.6 | 7.1 | 389 | 17 | US-10-242-535A-12486 | Sequence 12486, A |
| | 27 | 33.6 | 7.1 | 389 | 17 | US-10-085-783A-12486 | Sequence 12486, A |
| | 28 | 33.6 | 7.1 | 396 | 11 | US-09-969-034-578 | Sequence 578, App |
| | 29 | 33.6 | 7.1 | 447 | 10 | US-09-918-995-14121 | Sequence 14121, A |
| | 30 | 33.6 | 7.1 | 483 | 10 | US-09-918-995-13976 | Sequence 13976, A |
| | 31 | 33.6 | 7.1 | 674 | 16 | US-10-076-747-43 | Sequence 43, Appl |
| | 32 | 33.6 | 7.1 | 674 | 17 | US-10-001-885-78 | Sequence 78, Appl |
| | 33 | 33.6 | 7.1 | 676 | 11 | US-09-989-890-61 | Sequence 61, Appl |
| | 34 | 33.6 | 7.1 | 1344 | 17 | US-10-001-885-92 | Sequence 92, Appl |
| | 35 | 33.6 | 7.1 | 1375 | 17 | US-10-001-885-79 | Sequence 79, Appl |
| | 36 | 33.6 | 7.1 | 1756 | 13 | US-10-001-887-81 | Sequence 81, Appl |
| | 37 | 33.6 | 7.1 | 2731748 | 18 | US-10-297-465A-1 | Sequence 1, Appl |
| | 38 | 33 | 7.0 | 149382 | 19 | US-10-741-600-17661 | Sequence 17661, A |
| | 39 | 32.8 | 7.0 | 4251 | 16 | US-10-240-865-106 | Sequence 106, App |
| | 40 | 32.8 | 7.0 | 41694 | 18 | US-10-719-993-6806 | Sequence 6806, App |
| | 41 | 32.8 | 7.0 | 194883 | 13 | US-10-087-192-826 | Sequence 826, App |
| | 42 | 32.2 | 6.8 | 657 | 18 | US-10-767-701-2873 | Sequence 2873, App |
| | 43 | 32.2 | 6.8 | 1119 | 17 | US-10-242-355-195 | Sequence 195, App |
| | 44 | 32.2 | 6.8 | 1738 | 18 | US-10-739-930-475 | Sequence 475, App |
| | 45 | 32.2 | 6.8 | 3148 | 18 | US-10-437-963-80145 | Sequence 80145, A |

ALIGNMENTS

RESULT 1

US-09-785-548-3

/ Sequence 3, Application US/09785548

/ Patent No. US2002015577A1

/ GENERAL INFORMATION:

/ APPLICANT: AVENTIS PHARMACEUTICALS, INC.

/ TITLE OF INVENTION: COMPOSITIONS THAT CAN BE USED FOR REGULATING THE ACTIVITY OF PARK

/ FILE REFERENCE: ST000005

/ CURRENT APPLICATION NUMBER: US/09/785,548

/ CURRENT FILING DATE: 2001-02-20

/ NUMBER OF SEQ ID NOS: 46

/ SOFTWARE: PatentIn Ver. 2.1

/ SEQ ID NO 3

/ LENGTH: 471

/ TYPE: DNA

/ ORGANISM: Homo sapiens

/ FEATURE:

/ NAME/KEY: CDS

/ LOCATION: (1)..(471)

US-09-785-548-3

Query Match 100.0%; Score 471; DB 9; Length 471;

Best Local Similarity 100.0%; Pred. No. 9.5e-157; Indels 0; Gaps 0;

Matches 471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| Qy | 1 | GGAAGTCACAGGATGATCAATCTACAAAGCTTTATGTGTATTTGCAAAAGCCCTGT | 60 |
| | | | |
| Db | 1 | GGAAGTCACAGGATGATCAATCTACAAAGCTTTATGTGTATTTGCAAAAGCCCTGT | 60 |
| | | | |
| Qy | 61 | CNAAGATGTCAGCCCGGAAAACTCAGGATACAGTCAGCAGCAGCAGCAAGCTTC | 120 |
| | | | |
| Db | 61 | CNAAGATGTCAGCCCGGAAAACTCAGGATACAGTCAGCAGCAGCAGCAAGCTTC | 120 |
| | | | |
| Qy | 121 | ACCTTGACCGAGGATCATCTTGGCTGGATGATGTTTAAATTCAAACCGATGAGTGT | 180 |

```
Db 121 ACCTGACCCGAGGCTCATCTGCTGGAGTATGTTTAAATCCAAACGGATAGTGT 180
Oy 181 GAATGCCAATCCCACTGCTCCCTGGAGCTAGTGCAGAAATTTTAAATGTGAGCA 240
Db 181 GAATGCCAATCCCACTGCTCCCTGGAGCTAGTGCAGAAATTTTAAATGTGAGCA 240
Oy 241 CACCCCACTCTGACAAAGAAACATCAGTAGCTTTGCACTGATGCAACAAATAGTCG 240
Db 241 CACCCCACTCTGACAAAGAAACATCAGTAGCTTTGCACTGATGCAACAAATAGTCG 240
Oy 301 AACATCACTTGCATTAACGTGCAAGAGCTGAGAGCCCGCTGCTTCCAGTGCAC 360
Db 301 AACATCACTTGCATTAACGTGCAAGAGCTGAGAGCCCGCTGCTTCCAGTGCAC 360
Oy 361 TCCCGCACGTGATTTGCTTAGACTGTTTCCACTTAACGTGTGCAAGACTCAATGAT 420
Db 361 TCCCGCACGTGATTTGCTTAGACTGTTTCCACTTAACGTGTGCAAGACTCAATGAT 420
Oy 421 CGGAGTTTGTTCAGAACCCCTCAACTTGCTACTCCCTGCTGTGTGTAG 471
Db 421 CGGAGTTTGTTCAGAACCCCTCAACTTGCTACTCCCTGCTGTGTGTAG 471
```

RESULT 2
US-10-622-817-10

```
; Sequence 10, Application US/10622817
; Publication No. US20040214763A1
; GENERAL INFORMATION:
; APPLICANT: CORTEI, Olga
; APPLICANT: HAMPE, Cornelia
; APPLICANT: BRICE, Alexis
; APPLICANT: POUDIER, Laurent
; APPLICANT: ROONEY, Thomas
; APPLICANT: FOURNIER, Alain
; TITLE OF INVENTION: METHOD FOR DETERMINING THE ABILITY OF A COMPOUND TO MODIFY THE
; FILE REFERENCE: FRAV2002-0020 US NP
; CURRENT APPLICATION NUMBER: US/0/622,817
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: GB 0229934.5
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(471)
US-10-622-817-10
```

Query Match 100.0%; Score 471; DB 18; Length 471;
Best Local Similarity 100.0%; Pred. No. 9.5e-157;
Matches 471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Oy 1 GGAATGCCAGAGTAGATCAATCTAACAAGCTTTTATGTGATGCAAAAGCCCTGT 60
Db 1 GGAATGCCAGAGTAGATCAATCTAACAAGCTTTTATGTGATGCAAAAGCCCTGT 60
Oy 61 CAAAGAGTGCAGCGGGAACCACTCAGGTACAGTGCAGCACTGCAAGCAAGCGTC 120
Db 61 CAAAGAGTGCAGCGGGAACCACTCAGGTACAGTGCAGCACTGCAAGCAAGCGTC 120
Oy 121 ACCTGACCCAGGCTCATCTGCTGGAGTATGTTTAAATTTCCAAACCGGATAGTGT 180
Db 121 ACCTGACCCAGGCTCATCTGCTGGAGTATGTTTAAATTTCCAAACCGGATAGTGT 180
Oy 181 GAATGCCAATCCCACTGCTCCCTGGAGCTAGTGCAGAAATTTTAAATGTGAGCA 240
Db 181 GAATGCCAATCCCACTGCTCCCTGGAGCTAGTGCAGAAATTTTAAATGTGAGCA 240
```

```
Db 181 GAATGCCAATCCCACTGCTCCCTGGAGCTAGTGCAGAAATTTTAAATGTGAGCA 240
Oy 241 CACCCCACTCTGACAAAGAAACATCAGTAGCTTTGCACTGATGCAACAAATAGTCG 300
Db 241 CACCCCACTCTGACAAAGAAACATCAGTAGCTTTGCACTGATGCAACAAATAGTCG 300
Oy 301 AACATCACTTGCATTAACGTGCAAGAGCTGAGAGCCCGCTGCTTCCAGTGCAC 360
Db 301 AACATCACTTGCATTAACGTGCAAGAGCTGAGAGCCCGCTGCTTCCAGTGCAC 360
Oy 361 TCCCGCACGTGATTTGCTTAGACTGTTTCCACTTAACGTGTGCAAGACTCAATGAT 420
Db 361 TCCCGCACGTGATTTGCTTAGACTGTTTCCACTTAACGTGTGCAAGACTCAATGAT 420
Oy 421 CGGAGTTTGTTCAGAACCCCTCAACTTGCTACTCCCTGCTGTGTGTAG 471
Db 421 CGGAGTTTGTTCAGAACCCCTCAACTTGCTACTCCCTGCTGTGTGTAG 471
```

RESULT 3
US-10-839-688-12

```
; Sequence 12, Application US/10839688
; Publication No. US20050014173A1
; GENERAL INFORMATION:
; APPLICANT: Farret, Matthew J.
; TITLE OF INVENTION: PARKINSON'S DISEASE MARKERS
; FILE REFERENCE: 07039-448001
; CURRENT APPLICATION NUMBER: US/10/839,688
; PRIOR FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: US 60/468,832
; NUMBER OF SEQ ID NOS: 08
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-839-688-12
```

Query Match 99.0%; Score 466.4; DB 19; Length 2955;
Best Local Similarity 99.8%; Pred. No. 1.1e-154;
Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Oy 1 GGAATGCCAGAGTAGATCAATCTAACAAGCTTTTATGTGATGCAAAAGCCCTGT 60
Db 499 GGAATGCCAGAGTAGATCAATCTAACAAGCTTTTATGTGATGCAAAAGCCCTGT 558
Oy 61 CAAAGAGTGCAGCGGGAACCACTCAGGTACAGTGCAGCACTGCAAGCAAGCGTC 120
Db 559 CAAAGAGTGCAGCGGGAACCACTCAGGTACAGTGCAGCACTGCAAGCAAGCGTC 618
Oy 121 ACCTGACCCAGGCTCATCTGCTGGAGTATGTTTAAATTTCCAAACCGGATAGTGT 180
Db 619 ACCTGACCCAGGCTCATCTGCTGGAGTATGTTTAAATTTCCAAACCGGATAGTGT 678
Oy 181 GAATGCCAATCCCACTGCTCCCTGGAGCTAGTGCAGAAATTTTAAATGTGAGCA 240
Db 679 GAATGCCAATCCCACTGCTCCCTGGAGCTAGTGCAGAAATTTTAAATGTGAGCA 738
Oy 241 CACCCCACTCTGACAAAGAAACATCAGTAGCTTTGCACTGATGCAACAAATAGTCG 300
Db 739 CACCCCACTCTGACAAAGAAACATCAGTAGCTTTGCACTGATGCAACAAATAGTCG 798
Oy 301 AACATCACTTGCATTAACGTGCAAGAGCTGAGAGCCCGCTGCTTCCAGTGCAC 360
Db 799 AACATCACTTGCATTAACGTGCAAGAGCTGAGAGCCCGCTGCTTCCAGTGCAC 858
Oy 361 TCCCGCACGTGATTTGCTTAGACTGTTTCCACTTAACGTGTGCAAGACTCAATGAT 420
Db 859 TCCCGCACGTGATTTGCTTAGACTGTTTCCACTTAACGTGTGCAAGACTCAATGAT 420
Oy 421 CGGAGTTTGTTCAGAACCCCTCAACTTGCTACTCCCTGCTGTGTGTAG 471
Db 421 CGGAGTTTGTTCAGAACCCCTCAACTTGCTACTCCCTGCTGTGTGTAG 471
```

Db 919 CGGCGAGTTGTTCAGACGCTCACTTGCTACTCCCTGCTGTGTG 566

RESULT 4
US-10-473-226-1
; Sequence 1, Application US/10473226
; Publication No. US20040198650A1
; GENERAL INFORMATION:
; APPLICANT: NcGene A/S
; TITLE OF INVENTION: Means for inhibiting proteolytical processing of Parkin
; FILE REFERENCE: 506-204-WO
; CURRENT APPLICATION NUMBER: US/10/473,226
; PRIORITY FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: DK PA 2001 00525
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: US 60/281,286
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2960
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mRNA
; LOCATION: (1)..(2960)
; OTHER INFORMATION:
US-10-473-226-1

Query Match 99.0%; Score 466.4; DB 18; Length 2960;
Best Local Similarity 99.8%; Pred. No. 1,1e-154;
Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGAAGTCAGCAGGTAGATCATCTACACAGCTTTTATGTGTATGCAAGGCCCTGT 60
|
|
|
Db 504 GGAAGTCAGCAGGTAGATCATCTACACAGCTTTTATGTGTATGCAAGGCCCTGT 563
|
|
|
Qy 61 CAAAGATGACGCGGGAAAACTCAGGGTACAGTGCAGACCTGACAGGCGCAACGCTC 120
|
|
|
Db 564 CAAAGATGACGCGGGAAAACTCAGGGTACAGTGCAGACCTGACAGGCGCAACGCTC 623
|
|
|
Qy 121 ACCTTGACCAAGGATCCATCTTGTGCGATGATGATTTTAAATCCAAACCGGATAGTGT 180
|
|
|
Db 624 ACCTTGACCAAGGATCCATCTTGTGCGATGATGATTTTAAATCCAAACCGGATAGTGT 683
|
|
|
Qy 181 GAATGCCAATCCCACTGCGCTGCGAATGAGTGAATTTTCTTTAAATGTGAGCA 240
|
|
|
Db 684 GAATGCCAATCCCACTGCGCTGCGAATGAGTGAATTTTCTTTAAATGTGAGCA 743
|
|
|
Qy 241 CACCCCACTCTGACAGAAACATCACTAGCTTTGACCTGATCGCAAAATATGTCGG 300
|
|
|
Db 744 CACCCCACTCTGACAGAAACATCACTAGCTTTGACCTGATCGCAAAATATGTCGG 803
|
|
|
Qy 301 AACATCACTTGATTAAGTGCAGAGCTCAGAGCCCGCTCGTCTTTCCAGTGCAC 360
|
|
|
Db 804 AACATCACTTGATTAAGTGCAGAGCTCAGAGCCCGCTCGTCTTTCCAGTGCAC 863
|
|
|
Qy 361 TCCGCGACGTGATTTGCTTGAAGTCTTTCCATTATATGTTGACAAAGCTCAATGAT 420
|
|
|
Db 864 TCCGCGACGTGATTTGCTTGAAGTCTTTCCATTATATGTTGACAAAGCTCAATGAT 923
|
|
|
Qy 421 CGGAGTTTGTTCAGACGCTCAACTGGCTACTCCCTGCTGTGTG 468
|
|
|
Db 924 CGGAGTTTGTTCAGACGCTCAACTGGCTACTCCCTGCTGTGTG 971
|
|
|

RESULT 5
US-10-473-226-3
; Sequence 3, Application US/10473226
; Publication No. US20040198650A1
; GENERAL INFORMATION:
; APPLICANT: NcGene A/S
; TITLE OF INVENTION: Means for inhibiting proteolytical processing of Parkin

; FILE REFERENCE: 506-204-WO
; CURRENT APPLICATION NUMBER: US/10/473,226
; PRIORITY FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: DK PA 2001 00525
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: US 60/281,286
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2960
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mRNA
; LOCATION: (1)..(2960)
; OTHER INFORMATION:
US-10-473-226-3

Query Match 99.0%; Score 466.4; DB 18; Length 2960;
Best Local Similarity 99.8%; Pred. No. 1,1e-154;
Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGAAGTCAGCAGGTAGATCATCTACACAGCTTTTATGTGTATGCAAGGCCCTGT 60
|
|
|
Db 504 GGAAGTCAGCAGGTAGATCATCTACACAGCTTTTATGTGTATGCAAGGCCCTGT 563
|
|
|
Qy 61 CAAAGATGACGCGGGAAAACTCAGGGTACAGTGCAGACCTGACAGGCGCAACGCTC 120
|
|
|
Db 564 CAAAGATGACGCGGGAAAACTCAGGGTACAGTGCAGACCTGACAGGCGCAACGCTC 623
|
|
|
Qy 121 ACCTTGACCAAGGATCCATCTTGTGCGATGATGATTTTAAATCCAAACCGGATAGTGT 180
|
|
|
Db 624 ACCTTGACCAAGGATCCATCTTGTGCGATGATGATTTTAAATCCAAACCGGATAGTGT 683
|
|
|
Qy 181 GAATGCCAATCCCACTGCGCTGCGAATGAGTGAATTTTCTTTAAATGTGAGCA 240
|
|
|
Db 684 GAATGCCAATCCCACTGCGCTGCGAATGAGTGAATTTTCTTTAAATGTGAGCA 743
|
|
|
Qy 241 CACCCCACTCTGACAGAAACATCACTAGCTTTGACCTGATCGCAAAATATGTCGG 300
|
|
|
Db 744 CACCCCACTCTGACAGAAACATCACTAGCTTTGACCTGATCGCAAAATATGTCGG 803
|
|
|
Qy 301 AACATCACTTGATTAAGTGCAGAGCTCAGAGCCCGCTCGTCTTTCCAGTGCAC 360
|
|
|
Db 804 AACATCACTTGATTAAGTGCAGAGCTCAGAGCCCGCTCGTCTTTCCAGTGCAC 863
|
|
|
Qy 361 TCCGCGACGTGATTTGCTTGAAGTCTTTCCATTATATGTTGACAAAGCTCAATGAT 420
|
|
|
Db 864 TCCGCGACGTGATTTGCTTGAAGTCTTTCCATTATATGTTGACAAAGCTCAATGAT 923
|
|
|
Qy 421 CGGAGTTTGTTCAGACGCTCAACTGGCTACTCCCTGCTGTGTG 468
|
|
|
Db 924 CGGAGTTTGTTCAGACGCTCAACTGGCTACTCCCTGCTGTGTG 971
|
|
|

RESULT 6
US-10-622-817-7
; Sequence 7, Application US/10622817
; Publication No. US20040214763A1
; GENERAL INFORMATION:
; APPLICANT: CORTE, Olga
; APPLICANT: HAMPE, Cornelia
; APPLICANT: BRICE, Alexis
; APPLICANT: PRADIER, Laurent
; APPLICANT: ROONEY, Thomas
; APPLICANT: FOURNIER, Alain
; TITLE OF INVENTION: METHOD FOR DETERMINING THE ABILITY OF A COMPOUND TO MODIFY THE
; FILE REFERENCE: FRAV2002-0020 US NP
; INTERACTION BETWEEN PARKIN AND THE P38 PROTEIN
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: GB 0229934.5

;; PRIOR FILING DATE: 2002-12-20
;; PRIOR APPLICATION NUMBER: US 60/396,929
;; PRIOR FILING DATE: 2002-07-16
;; NUMBER OF SEQ ID NOS: 22
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 7
;; LENGTH: 2960
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (102)..(1499)
US-10-622-817-7

Query Match
Best Local Similarity 99.0%; Score 466.4; DB 18; Length 2960;
Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGTCCAGCAGGTGATCAATCTACACAGCTTTTATGTATTTGCAAGGCCCTGT 60
DB 504 GGAAGTCCAGCAGGTGATCAATCTACACAGCTTTTATGTATTTGCAAGGCCCTGT 60
QY 61 CAAAGAGTGCAGCCGGGAAAACCTCAGGTACAGTGCAGCAGCTGCAAGCAAGCTTC 120
DB 564 CAAAGAGTGCAGCCGGGAAAACCTCAGGTACAGTGCAGCAGCTGCAAGCAAGCTTC 120
QY 121 ACCTTGACCCAGGAGTCACTTGTGGGAGATGATGTTTAAATTCGAAACCGGATGATG 180
DB 624 ACCTTGACCCAGGAGTCACTTGTGGGAGATGATGTTTAAATTCGAAACCGGATGATG 180
QY 181 GAATGCCAATCCCACTGCTCCCTGGAGCTAGTGCAGAAATTTTCTTTAAATGTGAGCA 240
DB 684 GAATGCCAATCCCACTGCTCCCTGGAGCTAGTGCAGAAATTTTCTTTAAATGTGAGCA 240
QY 241 CACCCCACTCTGCAAGAAAACATCAGTAGCTTTGCACTGATGCCAACAAATAGTCG 300
DB 744 CACCCCACTCTGCAAGAAAACATCAGTAGCTTTGCACTGATGCCAACAAATAGTCG 300
QY 301 AACATCACTTGATTAAGTGCACAGCTCAGGAGCCGCTGCTGTTTCCAGTGCAAC 360
DB 804 AACATCACTTGATTAAGTGCACAGCTCAGGAGCCGCTGCTGTTTCCAGTGCAAC 360
QY 361 TCCCGCCAGCTGATTTGCTTGAAGTGTTCACCTTACTGTGTGACAAAGCTCAATGAT 420
DB 864 TCCCGCCAGCTGATTTGCTTGAAGTGTTCACCTTACTGTGTGACAAAGCTCAATGAT 420
QY 421 CGGCAATTGTTCAAGCACTGCACTTGAGTACTGCTGCTGCTGCTG 468
DB 924 CGGCAATTGTTCAAGCACTGCACTTGAGTACTGCTGCTGCTGCTG 971

RESULT 7
US-10-776-604-1
;; Sequence 1, Application US/10776604
;; Publication No. US2005000385A1
;; GENERAL INFORMATION:
;; APPLICANT: Shimizu, Nobuyoshi
;; APPLICANT: Mizuno, Yoshihumi
;; TITLE OF INVENTION: DNAs or Genes Participating in Parkinson's Disease
;; FILE REFERENCE: 0652.2110001
;; CURRENT FILING DATE: 2004-02-12
;; PRIOR APPLICATION NUMBER: US/10/776,604
;; PRIOR FILING DATE: 2000-08-09
;; PRIOR APPLICATION NUMBER: PCT/JP99/00545
;; PRIOR FILING DATE: 1999-02-09
;; PRIOR APPLICATION NUMBER: JP 10/27531
;; NUMBER OF SEQ ID NOS: 70
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 1
;; LENGTH: 2960
;; TYPE: DNA

;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (102)..(1496)
;; OTHER INFORMATION:
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (102)..(108)
;; OTHER INFORMATION: Exon 1
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (109)..(1272)
;; OTHER INFORMATION: Exon 2
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (273)..(513)
;; OTHER INFORMATION: Exon 3
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (514)..(635)
;; OTHER INFORMATION: Exon 4
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (636)..(719)
;; OTHER INFORMATION: Exon 5
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (720)..(835)
;; OTHER INFORMATION: Exon 6
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (836)..(972)
;; OTHER INFORMATION: Exon 7
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (973)..(1034)
;; OTHER INFORMATION: Exon 8
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1035)..(1184)
;; OTHER INFORMATION: Exon 9
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1185)..(1268)
;; OTHER INFORMATION: Exon 10
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1269)..(1386)
;; OTHER INFORMATION: Exon 11
;; NAME/KEY: misc feature
;; LOCATION: (1387)..(2960)
;; OTHER INFORMATION: Exon 12
US-10-776-604-1

Query Match
Best Local Similarity 99.0%; Score 466.4; DB 18; Length 2960;
Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGTCCAGCAGGTGATCAATCTACACAGCTTTTATGTATTTGCAAGGCCCTGT 60
DB 504 GGAAGTCCAGCAGGTGATCAATCTACACAGCTTTTATGTATTTGCAAGGCCCTGT 60
QY 61 CAAAGAGTGCAGCCGGGAAAACCTCAGGTACAGTGCAGCAGCTGCAAGCAAGCTTC 120
DB 564 CAAAGAGTGCAGCCGGGAAAACCTCAGGTACAGTGCAGCAGCTGCAAGCAAGCTTC 120
QY 121 ACCTTGACCCAGGAGTCACTTGTGGGAGATGATGTTTAAATTCGAAACCGGATGATG 180
DB 624 ACCTTGACCCAGGAGTCACTTGTGGGAGATGATGTTTAAATTCGAAACCGGATGATG 180
QY 181 GAATGCCAATCCCACTGCTCCCTGGAGCTAGTGCAGAAATTTTCTTTAAATGTGAGCA 240

Dh 684 GAATGCCAATCCCAACATGCTCCCTGGAGATGACGAGAAATTTTCTTAAATGTGAGCA 743
Qy 241 CACCCACCTCTGACAAAGAAACATCAGTCTTTGACCTTGATCGCAAAATAGTCGG 300
Db 744 CACCCACCTCTGACAAAGAAACATCAGTCTTTGACCTTGATCGCAAAATAGTCGG 803
Qy 301 AACATCACTTTCATTAAGTACAGACAGCTGAGAGCCCGCTCTGTTTCCAGTGCAC 360
Db 804 AACATCACTTTCATTAAGTACAGACAGCTGAGAGCCCGCTCTGTTTCCAGTGCAC 863
Qy 361 TCCCGCAGCGATTTGCTTACAGTCTTTTCCATTATCTGTGACAAAGCTCAATGAT 420
Db 864 TCCCGCAGCGATTTGCTTACAGTCTTTTCCATTATCTGTGACAAAGCTCAATGAT 923
Qy 421 CGGAGTTTGTTCAGACCTTCAACTGGCTACTCCCTGCTGTGTG 468
Db 924 CGGAGTTTGTTCAGACCTTCAACTGGCTACTCCCTGCTGTGTG 971

RESULT 8

US-10-839-688-11
; Sequence 11, Application US/10839688
; Publication No. US20050014173A1
; GENERAL INFORMATION:
; APPLICANT: Farret, Matthew J.
; TITLE OF INVENTION: PARKINSON'S DISEASE MARKERS
; FILE REFERENCE: 07039-448001
; CURRENT APPLICATION NUMBER: US/10/839, 688
; CURRENT FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: US 60/468,832
; PRIOR FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 2960
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-839-688-11

Query Match 99.0%; Score 466.4; DB 19; Length 2960;
Best Local Similarity 99.8%; Pred. No. 1.1e-154;
Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGAAGTCCAGAGGTAGATCAATCTACACAGCTTTTATGTGTAAGAGCCCTGT 60
Db 504 GGAAGTCCAGAGGTAGATCAATCTACACAGCTTTTATGTGTAAGAGCCCTGT 563
Qy 61 CAAAGATGACAGCCGGGAAAACACTCAGGGTACAGTGCACCTGACAGGCAACGCTC 120
Db 564 CAAAGATGACAGCCGGGAAAACACTCAGGGTACAGTGCACCTGACAGGCAACGCTC 623
Qy 121 ACCTTGACCCAGGGTCCATCTTCTGGGATGATGTTTAAATCCAAACCGGATGATG 180
Db 624 ACCTTGACCCAGGGTCCATCTTCTGGGATGATGTTTAAATCCAAACCGGATGATG 683
Qy 181 GAATGCCAATCCCAACCTGCGCTGGGACTAGTGCAGAAATTTTCTTAAATGTGAGCA 240
Db 684 GAATGCCAATCCCAACCTGCGCTGGGACTAGTGCAGAAATTTTCTTAAATGTGAGCA 743
Qy 241 CACCCACCTCTGACAAAGAAACATCAGTCTTTGACCTTGATCGCAAAATAGTCGG 300
Db 744 CACCCACCTCTGACAAAGAAACATCAGTCTTTGACCTTGATCGCAAAATAGTCGG 803
Qy 301 AACATCACTTTCATTAAGTACAGACAGCTGAGAGCCCGCTCTGTTTCCAGTGCAC 360
Db 804 AACATCACTTTCATTAAGTACAGACAGCTGAGAGCCCGCTCTGTTTCCAGTGCAC 863
Qy 361 TCCCGCAGCGATTTGCTTACAGTCTTTTCCATTATCTGTGACAAAGCTCAATGAT 420
Db 864 TCCCGCAGCGATTTGCTTACAGTCTTTTCCATTATCTGTGACAAAGCTCAATGAT 923
Qy 421 CGGAGTTTGTTCAGACCTTCAACTGGCTACTCCCTGCTGTGTG 468

Dh 924 CGGAGTTTGTTCAGACCTTCAACTGGCTACTCCCTGCTGTGTG 971

RESULT 9

US-10-622-817-13
; Sequence 13, Application US/10622817
; Publication No. US20040214763A1
; GENERAL INFORMATION:
; APPLICANT: CORTE, Olga
; APPLICANT: HAMPE, Cornelia
; APPLICANT: BRICE, Alexis
; APPLICANT: PRADIER, Laurent
; APPLICANT: ROONEY, Thomas
; APPLICANT: FOURNIER, Alain
; TITLE OF INVENTION: METHOD FOR DETERMINING THE ABILITY OF A COMPOUND TO MODIFY THE
; FILE REFERENCE: PRAV2002-0020 US NP
; CURRENT APPLICATION NUMBER: US/10/622,817
; PRIOR APPLICATION NUMBER: GB 0229934.5
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/396,929
; PRIOR FILING DATE: 2002-07-18
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 1470
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1470)
US-10-622-817-13

Query Match 73.7%; Score 347; DB 18; Length 1470;
Best Local Similarity 83.9%; Pred. No. 2.7e-112;
Matches 392; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Qy 2 GAAATCCAGAGGTAGATCAATCTACACAGCTTTTATGTGTAAGAGCCCTGT 61
Db 476 GAAATCCAGAGGTAGATCAATCTACACAGCTTTTATGTGTAAGAGCCCTGT 535
Qy 62 AAAGATGACAGCCGGGAAAACACTCAGGGTACAGTGCACACCTGACAGGCAACGCTCA 121
Db 536 AAAGATGACAGCCGGGAAAACACTCAGGGTACAGTGCACACCTGACAGGCAACGCTCA 595
Qy 122 CTTTGACCCAGGGTCCATCTTCTGGGATGATGTTTAAATCCAAACCGGATGATG 181
Db 596 CTTTGACCCAGGGTCCATCTTCTGGGATGATGTTTAAATCCAAACCGGATGATG 655
Qy 182 AATGCCAATCCCAACCTGCGCTGGGACTAGTGCAGAAATTTTCTTAAATGTGAGCAC 241
Db 656 AATGCCAATCCCAACCTGCGCTGGGACTAGTGCAGAAATTTTCTTAAATGTGAGCAC 715
Qy 242 ACCCACTCTGACAAAGAAACATCAGTCTTTGACCTTGATCGCAAAATAGTCGGA 301
Db 716 ACCCACTCTGACAAAGAAACATCAGTCTTTGACCTTGATCGCAAAATAGTCGGA 775
Qy 302 ACATCACTTTCATTAAGTACAGACAGCTGAGAGCCCGCTCTGTTTCCAGTGCAC 361
Db 776 ACATCACTTTCATTAAGTACAGACAGCTGAGAGCCCGCTCTGTTTCCAGTGCAC 835
Qy 362 CCGGCGAGGATTTGCTTACAGTCTTTTCCATTATCTGTGACAAAGCTCAATGATC 421
Db 836 CCGGCGAGGATTTGCTTACAGTCTTTTCCATTATCTGTGACAAAGCTCAATGATC 895
Qy 422 GGCAGTTTGTTCAGACCTTCAACTGGCTACTCCCTGCTGTGTG 468
Db 896 GGCAGTTTGTTCAGACCTTCAACTGGCTACTCCCTGCTGTGTG 942

RESULT 10
US-10-776-604-3

```
Sequence 3, Application US/10776604
Publication No. US2005000385A1
GENERAL INFORMATION:
APPLICANT: Shimizu, Nobuyoshi
APPLICANT: Mizuno, Yoshikumi
TITLE OF INVENTION: DNAs or Genes Participating in Parkinson's Disease
FILE REFERENCE: 0652.2110001
CURRENT APPLICATION NUMBER: US/10/776,604
CURRENT FILING DATE: 2004-02-12
PRIOR APPLICATION NUMBER: 09/601,844
PRIOR FILING DATE: 2000-08-09
PRIOR APPLICATION NUMBER: PCT/JP99/00545
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: JP 10/27531
PRIOR FILING DATE: 1998-02-09
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 2876
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (102)..(1412)
OTHER INFORMATION:
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102)..(108)
OTHER INFORMATION: Exon 1
FEATURE:
NAME/KEY: misc_feature
LOCATION: (109)..(272)
OTHER INFORMATION: Exon 2
FEATURE:
NAME/KEY: misc_feature
LOCATION: (273)..(513)
OTHER INFORMATION: Exon 3
FEATURE:
NAME/KEY: misc_feature
LOCATION: (514)..(635)
OTHER INFORMATION: Exon 4
FEATURE:
NAME/KEY: misc_feature
LOCATION: (636)..(751)
OTHER INFORMATION: Exon 6
FEATURE:
NAME/KEY: misc_feature
LOCATION: (752)..(888)
OTHER INFORMATION: Exon 7
FEATURE:
NAME/KEY: misc_feature
LOCATION: (889)..(950)
OTHER INFORMATION: Exon 8
FEATURE:
NAME/KEY: misc_feature
LOCATION: (951)..(1100)
OTHER INFORMATION: Exon 9
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1101)..(1184)
OTHER INFORMATION: Exon 10
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1185)..(1302)
OTHER INFORMATION: Exon 11
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1303)..(2876)
OTHER INFORMATION: Exon 12
US-10-776-604-3
```

```
Query Match 61.2%; Score 288.4; DB 18; Length 2876;
Best Local Similarity 81.8%; Pred. No. 2.9e-91;
```

```
Matches 383; Conservative 0; Mismatches 1; Indels 84; Gaps 1;
QY 1 GGAACTCCAGAGGTAGATCAATCTACAACAGCTTTATGTGATGCAAAAGCCCTGT 60
DB 504 GGAAGTCCAGCAGGTAGATCAATCTACAACAGCTTTATGTGATGCAAAAGCCCTGT 563
QY 61 CAAGAGTGCAGCCCGGAAAACTAGGGTACAGTGACACCTTGACGAGGCAACGCTC 120
DB 564 CAAGAGTGCAGCCCGGAAAACTAGGGTACAGTGACACCTTGACGAGGCAACGCTC 623
QY 121 ACCTTGACCCAGGGTTCATCTGTGGGATGATGTTTAATCCAAACCGAATGATGT 180
DB 624 ACCTTGACCCAGG----- 635
QY 181 GAATGCCAATCCCACTGCTGCGCTGGAGCTAGTGCAAGATTTTCTTTAATGTGAGCA 240
DB 636 -----GAATTTTCTTTAATGTGAGCA 659
QY 241 CACCCGACCTCTGACAGAAACATCAGTCTTTGACCTGATCCGACAAATAGTCGG 300
DB 660 CACCCGACCTCTGACAGAAACATCAGTCTTTGACCTGATCCGACAAATAGTCGG 719
QY 301 AACATCACTTGACATTAAGTGCACAGACGTCAGAGCCCGTCTGGTTTCCAGTCAAC 360
DB 720 AACATCACTTGACATTAAGTGCACAGACGTCAGAGCCCGTCTGGTTTCCAGTCAAC 779
QY 361 TCCCGCAGGTGATTTGCTTAAGTCTGTTTCACTTATCTGTGTGACAAAGACTCAATGAT 420
DB 780 TCCCGCAGGTGATTTGCTTAAGTCTGTTTCACTTATCTGTGTGACAAAGACTCAATGAT 839
QY 421 CGGAGTTTGTTCAGACACCTCAACTGGTACTCCCTGCTTGTTG 468
DB 840 CGGAGTTTGTTCAGACACCTCAACTGGTACTCCCTGCTTGTTG 887

RESULT 11
US-10-839-688-6
Sequence 6, Application US/10839688
Publication No. US20050014173A1
GENERAL INFORMATION:
APPLICANT: Farrer, Matthew J.
TITLE OF INVENTION: PARKINSON'S DISEASE MARKERS
FILE REFERENCE: 07039-448001
CURRENT APPLICATION NUMBER: US/10/839,688
CURRENT FILING DATE: 2004-05-05
PRIOR APPLICATION NUMBER: US 60/468,832
PRIOR FILING DATE: 2003-05-08
NUMBER OF SEQ ID NOS: 81
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 650
TYPE: DNA
ORGANISM: Homo sapiens
US-10-839-688-6

Query Match 29.4%; Score 138.6; DB 19; Length 650;
Best Local Similarity 99.3%; Pred. No. 3.4e-38;
Matches 138; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 330 CAGGAGCCCCGCTGCTGTTTCCAGTCACTCCGCGCAGGTGATTTGCTTAGACTGTTT 389
DB 208 CAGGAGCCCCGCTGCTGTTTCCAGTCACTCCGCGCAGGTGATTTGCTTAGACTGTTT 267
QY 390 CCACTTATAGTGTGACAAAGACTCAATGATCGGCAAGTTTGTTCAGCAGCCCTCAACTTG 449
DB 268 CCACTTATAGTGTGACAAAGACTCAATGATCGGCAAGTTTGTTCAGCAGCCCTCAACTTG 327
QY 450 CTACTCCCTGCTGTGTG 468
DB 328 CTACTCCCTGCTGTGTG 346

RESULT 12
```


| | | | | |
|----|-----|---|-----------------|-----|
| Db | 134 | CTTGACACGCTTTCATTTTCCTCCAAATGCTGCATGCCA | CACTTGAGTACGAAG | 193 |
| Qy | 263 | CATCAGTAGCTTTGACCTGATGCGACAAATAGTC | 298 | |
| Db | 194 | CANAAGTATTTTAAACATGACAGCTAANAACAATTC | 229 | |

Search completed: February 26, 2005, 13:19:16
Job time : 286.738 secs